

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 17:06:51 ; Search time 1507.88 Seconds
(without alignments)
7418.598 Million cell updates/sec

Title: US-09-671-050-9
Perfect score: 1041
Sequence: 1 atggaaaagtatgaaaaatt.....aggtaattccgctcaaaagt 1041

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 11351937 seqs, 5372889281 residues

Word size : 20

Total number of hits satisfying chosen parameters: 6380

inimum DB seq length: 0
aximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estom:*
6: em_estpl:*
7: em_estba:*
8: em_estro:*
9: em_estov:*
10: em_htc:*
11: gb_est1:*
12: gb_est2:*
13: gb_htc:*
14: gb_gss:*
15: em_gss_fun:*
16: em_gss_hum:*
17: em_gss_inv:*
18: em_gss_pln:*
19: em_gss_pro:*
20: em_gss_rod:*
21: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length DB ID	Description
C 1	198	19.0	295 10	AA626859 zu89f10.s
C 2	47	4.5	756 11	BG390099
C 3	41	3.9	433 11	BF869052 IL5-ET011
C 4	39	3.7	328 13	AQ230677 HS_2033_B
C 5	37	3.6	171 11	BI040145 RC5-NT026
C 6	37	3.6	368 13	AQ030066
C 7	37	3.6	494 13	B32077
C 8	37	3.6	700 13	AG014845 Homo sapi
C 9	36	3.5	207 10	BE150038 QV3-HT026
C 10	36	3.5	366 11	H95588
C 11	36	3.5	408 10	AA678313 z116c01.s1
C 12	36	3.5	500 13	AQ769170 HS_3155_B

C 13	36	3.5	516 13	AQ668860 HS_5431_A
C 14	36	3.5	801 11	BG698963 602703384
C 15	35	3.4	193 10	BE149958 RC0-HT025
C 16	35	3.4	501 13	AQ617734 HS_5161_B
C 17	35	3.4	720 11	BF964650 602267728
C 18	35	3.4	778 13	AQ321133 RPC111-11
C 19	34	3.3	369 11	BF878219 CM2-ET012
C 20	34	3.3	377 13	AQ346239 RPC111-11
C 21	34	3.3	407 10	AW864431 PM4-SN001
C 22	34	3.3	412 13	AQ093501 HS_3018_B
C 23	34	3.3	589 10	AW958514 EST370584
C 24	34	3.3	734 10	BE296260 601174838
C 25	33	3.2	279 10	AI051456 ow29b12.s
C 26	33	3.2	298 10	AW339622 he15c05.x
C 27	33	3.2	348 10	AI244758 qj97g12.x
C 28	33	3.2	400 11	R98160 Yq66h03.r1
C 29	33	3.2	405 13	AQ142429 HS_2244_A
C 30	33	3.2	409 13	AQ455559 HS_5157_B
C 31	33	3.2	411 11	M78087 EST01677 Su
C 32	33	3.2	417 10	AA457230 aa91c02.s
C 33	33	3.2	447 10	BE146457 QV0-HT021
C 34	33	3.2	450 11	BF802026 CM1-CI009
C 35	33	3.2	453 10	AA984191 am82b06.s
C 36	33	3.2	679 11	BF982842 602305602
C 37	33	3.2	712 13	AQ595304 HS_5448_A
C 38	33	3.2	859 11	BG575508 602598919
C 39	33	3.2	874 11	BG912455 602806793
C 40	32	3.1	235 11	BG151645 nai68c04.
C 41	32	3.1	392 13	AQ095034 HS_3028_A
C 42	32	3.1	616 13	AG002404 Homo sapi
C 43	32	3.1	658 13	AG002403 Homo sapi
C 44	32	3.1	678 13	AG002384 Homo sapi
C 45	32	3.1	687 10	AL048065 DKFzp586G
C 46	32	3.1	691 13	AG002385 Homo sapi
C 47	32	3.1	853 13	AQ742378 HS_5382_B
C 48	31	3.0	248 10	AA640211 nn95h08.s
C 49	31	3.0	349 13	AQ341086 RPC111-11
C 50	31	3.0	439 10	AI814697 wk66d08.x
C 51	31	3.0	475 10	AL120748 DKFzp762M
C 52	31	3.0	589 10	AL134562 DKFzp547B
C 53	31	3.0	705 11	BG701158 602680728
C 54	31	3.0	837 11	BG400082 602442158
C 55	31	3.0	942 10	BE547155 601072573
C 56	31	3.0	945 11	BF966168 602286457
C 57	30	2.9	145 11	N74894 za88c04.s1
C 58	30	2.9	229 10	AA503246 ne45g08.s
C 59	30	2.9	237 10	AA302327 EST15447
C 60	30	2.9	255 10	AI567388 tp67d03.x
C 61	30	2.9	258 10	AA525875 ni57e09.s
C 62	30	2.9	288 11	BI007270 MRI-RF007
C 63	30	2.9	335 10	AA232416 zr23d03.r
C 64	30	2.9	336 10	AI287787 qv68f05.x
C 65	30	2.9	339 13	B43053 HS-1057-A1-
C 66	30	2.9	345 10	BE089717 QV0-BF070
C 67	30	2.9	352 10	AI254242 qt90f03.x
C 68	30	2.9	375 10	AA831906 oe21f04.s
C 69	30	2.9	378 10	AI204643 qd44a04.x
C 70	30	2.9	394 10	AA826078 oe24b03.s
C 71	30	2.9	405 10	AA666327 ac43c08.s
C 72	30	2.9	434 10	AW168900 xj14h10.x
C 73	30	2.9	435 10	AW675053 ba60g11.y
C 74	30	2.9	440 11	N66948 za48e10.s1
C 75	30	2.9	449 10	AI927393 wo75b10.x
C 76	30	2.9	452 10	AA626446 ab49h03.r
C 77	30	2.9	454 10	AI475902 tc95f02.x
C 78	30	2.9	455 10	AA157881 zo69g09.s
C 79	30	2.9	456 10	AI753334 cr08f09.x
C 80	30	2.9	462 10	AW972786 EST384881
C 81	30	2.9	478 13	AQ756485 HS_5354_B
C 82	30	2.9	481 11	N70325 za59h12.s1
C 83	30	2.9	511 10	AA601240 no14g03.s
C 84	30	2.9	526 13	AQ190209 HS_3218_A
C 85	30	2.9	540 13	B70510 CIT-HSP-205

c 86	30	2.9	542	10	AI344320	AI344320 tc03e11.x	159	28	2.7	279	10	AA302660	AA302660 EST10167
c 87	30	2.9	557	10	AV717771	AV717771 AV717771	160	28	2.7	279	10	AA605266	AA605266 no83c03.s
c 88	30	2.9	562	11	W05737	W05737 za88c04.r1	c 161	28	2.7	286	10	AA825946	AA825946 oe23f05.s
c 89	30	2.9	582	10	BE155132	BE155132 PM1-HT035	162	28	2.7	286	10	AA834766	AA834766 oe01d05.s
c 90	30	2.9	604	13	AQ423585	AQ423585 CITBI-EI-	c 163	28	2.7	286	13	AQ072686	AQ072686 HS_2249_B
c 91	30	2.9	637	13	AQ936375	AQ936375 RPCI-11-S	c 164	28	2.7	288	11	F00153	F00153 HSB37E011.S
c 92	30	2.9	703	11	BG706072	BG706072 60269027	c 165	28	2.7	291	10	AA719769	AA719769 zh38f04.s
c 93	30	2.9	779	13	AQ739207	AQ739207 HS_5382_B	166	28	2.7	293	10	AW104788	AW104788 xd45c03.x
c 94	30	2.9	810	10	AL519355	AL519355 AL519355	167	28	2.7	295	11	BG978475	BG978475 CM2-CN006
c 95	30	2.9	867	10	AL519354	AL519354 AL519354	c 168	28	2.7	295	11	F31605	F31605 HSPD22913.H
c 96	30	2.9	868	10	BE544285	BE544285 601076760	c 169	28	2.7	296	13	AQ111420	AQ111420 CIT-HSP-2
c 97	30	2.9	1011	11	BF340406	BF340406 602036759	c 170	28	2.7	301	13	AQ095016	AQ095016 HS_3028_A
c 98	30	2.9	1020	11	BG615396	BG615396 602643974	c 171	28	2.7	311	13	AQ422892	AQ422892 CITBI-EI-
c 99	29	2.8	234	13	AQ200492	AQ200492 RPCI11-63	c 172	28	2.7	312	10	AA228727	AA228727 nc46c05.r
c 100	29	2.8	249	11	BI005553	BI005553 CM0-HN020	c 173	28	2.7	313	10	AA846482	AA846482 ai85e08.s
c 101	29	2.8	261	11	BF877322	BF877322 QV2-ET010	174	28	2.7	318	10	AW517900	AW517900 xt66a04.x
c 102	29	2.8	318	10	AI432232	AI432232 th43c07.x	175	28	2.7	323	13	AQ487385	AQ487385 RPCI-11-2
c 103	29	2.8	352	10	AA515254	AA515254 ng69e10.s	176	28	2.7	324	11	H38416	H38416 yn82g05.sl
c 104	29	2.8	377	11	H58407	H58407 yr20g12.r1	c 177	28	2.7	325	10	AA584183	AA584183 no11a07.s
c 105	29	2.8	391	10	AI865092	AI865092 wk07d06.x	c 178	28	2.7	326	10	AA886560	AA886560 ny42a09.s
c 106	29	2.8	392	13	AQ567558	AQ567558 HS_2118_B	c 179	28	2.7	327	10	AA366543	AA366543 EST77510
c 107	29	2.8	393	13	AQ114347	AQ114347 CIT-HSP-2	c 180	28	2.7	328	11	BG960534	BG960534 PM2-CT080
c 108	29	2.8	397	11	R33826	R33826 Yh78g09.sl	181	28	2.7	330	13	AQ504662	AQ504662 RPCI-11-2
c 109	29	2.8	407	10	AA122219	AA122219 zn29c03.s	182	28	2.7	331	10	AI472766	AI472766 tal3f05.x
c 110	29	2.8	417	10	AW021007	AW021007 df17f06.y	c 183	28	2.7	333	11	BF763423	BF763423 CM1-CS003
c 111	29	2.8	437	10	BE065900	BE065900 RC2-BT031	c 184	28	2.7	333	13	AQ581730	AQ581730 RPCI-11-4
c 112	29	2.8	448	10	AA837335	AA837335 Q68h02.s	c 185	28	2.7	339	13	AQ047180	AQ047180 RPCI11-35
c 113	29	2.8	459	10	BE065895	BE065895 RC2-BT031	c 186	28	2.7	348	10	AI344962	AI344962 tb01d09.x
c 114	29	2.8	463	10	BE065796	BE065796 RC2-BT031	c 187	28	2.7	351	10	AV652978	AV652978 AV652978
c 115	29	2.8	469	11	BF229305	BF229305 PM4-AN009	188	28	2.7	353	10	AA347203	AA347203 EST53463
c 116	29	2.8	469	11	BF229310	BF229310 PM4-AN009	189	28	2.7	354	11	HA4786	HA4786 yp20g12.r1
c 117	29	2.8	470	13	B80116	B80116 CIT-HSP-204	190	28	2.7	356	10	AA304577	AA304577 EST175803
c 118	29	2.8	475	13	B63124	B63124 CIT978SK-A-	191	28	2.7	358	10	BE175609	BE175609 RC5-HT058
c 119	29	2.8	477	10	BE065817	BE065817 RC2-BT031	c 192	28	2.7	358	13	AQ451866	AQ451866 HS_5183_A
c 120	29	2.8	491	11	T77730	T77730 yd16a11.sl	c 193	28	2.7	359	10	BE042026	BE042026 hj13f12.y
c 121	29	2.8	497	10	AW103196	AW103196 xg43f09.x	c 194	28	2.7	361	10	BE043735	BE043735 hk42g06.y
c 122	29	2.8	500	13	AQ548911	AQ548911 RPCI-11-4	c 195	28	2.7	361	10	BE175589	BE175589 RC5-HT058
c 123	29	2.8	529	13	AQ118113	AQ118113 HS_3001_A	c 196	28	2.7	362	10	AA302964	AA302964 EST113120
c 124	29	2.8	541	11	BF887415	BF887415 QV3-TN016	c 197	28	2.7	364	10	BE149242	BE149242 RC2-HT025
c 125	29	2.8	558	13	B68229	B68229 CIT978SK-A-	198	28	2.7	366	10	AA577815	AA577815 nn24d05.s
c 126	29	2.8	560	13	AQ937863	AQ937863 NB6-591C	199	28	2.7	366	13	AQ541818	AQ541818 RPCI-11-3
c 127	29	2.8	570	10	AI742503	AI742503 wg43f11.x	c 200	28	2.7	368	11	H00437	H00437 YJ24a02.sl
c 128	29	2.8	572	10	AW939304	AW939304 QV1-DT007	c 201	28	2.7	370	13	AQ695877	AQ695877 HS_5491_B
c 129	29	2.8	574	10	AL042377	AL042377 DXF2p434A	c 202	28	2.7	377	11	BG988097	BG988097 MR2-HT116
c 130	29	2.8	590	13	AQ543668	AQ543668 RPCI-11-3	c 203	28	2.7	379	10	AI760985	AI760985 wh97a10.x
c 131	29	2.8	602	10	BE147684	BE147684 RC3-HT023	c 204	28	2.7	381	13	AQ027138	AQ027138 CIT-HSP-2
c 132	29	2.8	649	13	AQ544103	AQ544103 RPCI-11-3	c 205	28	2.7	384	13	AQ199309	AQ199309 RPCI11-58
c 133	29	2.8	670	11	BF832615	BF832615 PM0-HT091	c 206	28	2.7	385	11	BG001009	BG001009 QV0-GN021
c 134	29	2.8	767	13	AQ742589	AQ742589 HS_5384_B	c 207	28	2.7	386	11	BG005228	BG005228 CM2-GN022
c 135	28	2.7	113	10	AA728909	AA728909 nw01a04.s	c 208	28	2.7	391	13	AQ234691	AQ234691 HS_2045_A
c 136	28	2.7	128	13	AQ471829	AQ471829 CITBI-EI-	c 209	28	2.7	394	13	B46875	B46875 HS-1066-A1-
c 137	28	2.7	129	10	AI478685	AI478685 tm01b07.x	210	28	2.7	395	13	AQ541822	AQ541822 RPCI-11-3
c 138	28	2.7	136	11	BI000436	BI000436 MR3-HN006	c 211	28	2.7	397	10	AI348892	AI348892 tb05d09.x
c 139	28	2.7	150	10	AA491440	AA491440 ab01a12.s	c 212	28	2.7	397	11	BG005129	BG005129 CM2-GN022
c 140	28	2.7	170	11	BG995538	BG995538 CM0-HT129	c 213	28	2.7	400	13	AQ771299	AQ771299 HS_2106_A
c 141	28	2.7	190	10	AA346883	AA346883 EST53099	c 214	28	2.7	401	10	AI130838	AI130838 qc09h11.x
c 142	28	2.7	205	13	B81851	B81851 RPCI11-17L4	c 215	28	2.7	402	10	AA548438	AA548438 nk03a11.s
c 143	28	2.7	208	10	AI266021	AI266021 qp63b10.x	c 216	28	2.7	410	11	R19998	R19998 Yg38h06.r1
c 144	28	2.7	208	10	AW301555	AW301555 xs78f07.x	c 217	28	2.7	414	11	N27362	N27362 Yw52g05.sl
c 145	28	2.7	209	11	BF997387	BF997387 QV3-GN020	c 218	28	2.7	415	13	AQ529967	AQ529967 RPCI-11-3
c 146	28	2.7	219	10	AA910123	AA910123 oe28h02.s	c 219	28	2.7	416	10	AA846966	AA846966 oe12d07.s
c 147	28	2.7	219	13	AZ693711	AZ693711 AST-2HLB4	c 220	28	2.7	416	10	BE155099	BE155099 PM1-HT035
c 148	28	2.7	223	10	AA828919	AA828919 od81a04.s	c 221	28	2.7	416	13	AQ481593	AQ481593 RPCI-11-2
c 149	28	2.7	231	10	AA298532	AA298532 EST114149	c 222	28	2.7	417	11	BG957324	BG957324 IL3-CT067
c 150	28	2.7	234	10	AA296569	AA296569 EST11094	c 223	28	2.7	420	11	R18066	R18066 Yg15e11.r1
c 151	28	2.7	242	11	BF343920	BF343920 602015607	224	28	2.7	422	10	AA729755	AA729755 nx61d06.s
c 152	28	2.7	253	13	B53923	B53923 CIT-HSP-201	225	28	2.7	422	11	BF887129	BF887129 PM4-TN017
c 153	28	2.7	255	11	F19609	F19609 HSPD04214.H	c 226	28	2.7	422	11	BF897648	BF897648 RC1-MT016
c 154	28	2.7	265	13	AQ508343	AQ508343 RPCI-11-2	c 227	28	2.7	422	11	BF929547	BF929547 QV3-NT021
c 155	28	2.7	269	11	T20109	T20109 B466R Heart	c 228	28	2.7	426	10	AA551962	AA551962 nj83d08.s
c 156	28	2.7	273	13	AQ667078	AQ667078 HS_2107_A	c 229	28	2.7	426	13	AQ552478	AQ552478 RPCI-11-3
c 157	28	2.7	275	11	T69571	T69571 YC44D09.sl	230	28	2.7	429	11	BF836915	BF836915 CM3-HT099
c 158	28	2.7	275	11	BF847690	BF847690 IL5-EN008	c 231	28	2.7	431	10	AV763196	AV763196 AV763196

159	28	2.7	279	10	AA302660	AA302660
160	28	2.7	279	10	AA605266	AA605266
c 161	28	2.7	286	10	AA825946	AA825946
162	28	2.7	286	10	AA834766	AA834766
c 163	28	2.7	286	13	AQ072686	AQ072686
c 164	28	2.7	288	11	F00153	F00153
165	28	2.7	291	10	AA719769	AA719769
166	28	2.7	293	10	AW104788	AW104788
167	28	2.7	295	11	BG978475	BG978475
c 168	28	2.7	295	11	F31605	F31605
c 169	28	2.7	296	13	AQ111420	AQ111420
170	28	2.7	301	13	AQ095016	AQ095016
c 171	28	2.7	311	13	AQ422892	AQ422892
c 172	28	2.7	312	10	AA228727	AA228727
173	28	2.7	313	10	AA846482	AA846482
174	28	2.7	318	10	AW517900	AW517900
175	28	2.7	323	13	AQ487385	AQ487385
176	28	2.7	324	11	H38416	H38416
c 177	28	2.7	325	10	AA584183	AA584183
c 178	28	2.7	326	10	AA886560	AA886560
c 179	28	2.7	327	10	AA366543	AA366543
c 180	28	2.7	328	11	BG960534	BG960534
181	28	2.7	330	13	AQ504662	AQ504662
182	28	2.7	331	10	AI472766	AI472766
c 183	28	2.7	333	11	BF763423	BF763423
c 184	28	2.7	333	13	AQ581730	AQ581730
185	28	2.7	339	13	AQ047180	AQ047180
c 186	28	2.7	348	10	AI344962	AI344962
c 187	28	2.7	351	10	AV652978	AV652978
c 188	28	2.7	353	10	AA347203	AA347203
189	28	2.7	354	11	H44786	H44786
190	28	2.7	356	10	AA304577	AA304577
191	28	2.7	358	10	BE175609	BE175609
c 192	28	2.7	358	13	AQ451866	AQ451866
c 193	28	2.7	359	10	BE042026	BE042026
c 194	28	2.7	361	10	BE043735	BE043735
195	28	2.7	361	10	BE175589	BE175589
196	28	2.7	362	10	AA302964	AA302964
c 197	28	2.7	364	10	BE149242	BE149242
198	28	2.7	366	10	AA577815	AA577815
199	28	2.7	366	13	AQ541818	AQ541818
200	28	2.7	368	11	H00437	H00437
c 201	28	2.7	370	13	AQ695877	AQ695877
c 202	28	2.7	377	11	BG988097	BG988097
203	28	2.7	379	10	AI760985	AI760985
c 204	28	2.7	381	13	AQ027138	AQ027138
205	28	2.7	384	13	AQ199309	AQ199309
c 206	28	2.7	385	11	BG001009	BG001009
c 207	28	2.7	386	11	BG005228	BG005228
c 208	28	2.7	391	13	AQ234691	AQ234691
209	28	2.7	394	13	B46875	B46875
210	28	2.7	395	13	AQ541822	AQ541822
c 211	28	2.7	397	10	AI348892	AI348892
c 212	28	2.7	397	11	BG005129	BG005129
c 213	28	2.7	400	13	AQ771299	AQ771299
214	28	2.7	401	10	AI130838	AI130838
c 215	28	2.7	402	10	AA548438	AA548438
c 216	28	2.7	410	11	R19998	R19998
217	28	2.7	414	11	N27362	N27362
218	28	2.7	415	13	AQ529967	AQ529967
c 219	28	2.7	416	10	AA846966	AA846966
c 220	28	2.7	416	10	BE155099	BE155099
221	28	2.7	416	13	AQ481593	AQ481593
c 222	28	2.7	417	11	BG957324	BG957324
223	28	2.7	420	11	R18066	R18066
224	28	2.7	422	10	AA729755	AA729755
225	28	2.7	422	11	BF887129	BF887129
c 226	28	2.7	422	11	BF897648	BF897648
227	28	2.7	422	11	BF929547	BF929547
c 228	28	2.7	426	10	AA551962	AA551962
229	28	2.7	426	13	AQ552478	AQ552478
230	28	2.7	429	11	BF836915	BF836915
c 231	28	2.7	431	10	AV763196	AV763196

232	2.7	432	11	H46988
233	2.7	433	10	AW277123
234	2.7	434	13	AQ043528
235	2.7	439	10	AV732618
c	2.7	439	11	BF798467
	2.7	440	11	N22058
237	2.7	441	10	AA325466
238	2.7	444	10	AI821493
c	2.7	444	10	BE000907
	2.7	448	11	H19684
241	2.7	450	10	AI423231
242	2.7	450	13	AQ267366
243	2.7	451	10	AI732074
c	2.7	458	11	BE887225
	2.7	458	11	BF887098
246	2.7	458	11	BF887098
247	2.7	459	10	AI820988
248	2.7	460	10	AA603502
249	2.7	460	11	R92256
c	2.7	460	13	B72538
	2.7	462	10	AA015971
251	2.7	462	10	BE150018
252	2.7	462	11	BF914803
253	2.7	462	11	AI338899
254	2.7	464	10	BE138593
255	2.7	465	10	BE138593
c	2.7	465	13	AQ178738
	2.7	466	11	BF926463
257	2.7	467	13	AQ167841
c	2.7	469	10	AA525927
	2.7	475	11	BF770393
260	2.7	475	13	AQ783476
c	2.7	484	11	BG285245
	2.7	485	13	AQ550476
264	2.7	486	11	BI003340
265	2.7	488	11	BG619838
c	2.7	490	11	BF843512
	2.7	491	10	AL044807
267	2.7	493	13	B36003
268	2.7	495	10	AI472093
269	2.7	498	10	AL044445
270	2.7	500	13	AQ471002
c	2.7	501	10	AA074774
	2.7	501	13	AQ261648
273	2.7	514	13	AQ175560
274	2.7	517	10	AI379712
c	2.7	521	10	AW936276
	2.7	521	13	AQ122627
277	2.7	523	11	BI029851
278	2.7	523	13	B69012
279	2.7	527	13	AQ184178
280	2.7	528	13	AQ321460
c	2.7	531	10	AU158549
	2.7	539	13	AQ783497
282	2.7	542	11	BF821018
283	2.7	542	13	AQ483630
284	2.7	543	13	B15366
c	2.7	544	11	BF737356
	2.7	544	13	AQ213301
285	2.7	545	13	AQ491848
c	2.7	546	13	AQ410315
	2.7	547	10	AA147076
289	2.7	549	10	BE006648
290	2.7	550	13	AQ701930
291	2.7	556	11	BF812313
292	2.7	557	11	N63823
293	2.7	557	13	AQ496544
294	2.7	562	13	AQ456563
c	2.7	570	13	AQ311678
	2.7	579	10	AA206366
300	2.7	579	11	BF340510
301	2.7	580	11	BF925837
302	2.7	582	10	AA577696
303	2.7	582	10	AA577696
304	2.7	582	10	AA577696

c	305	28	2.7	582	13	AQ264917
	306	28	2.7	584	11	BI010689
	307	28	2.7	586	13	AQ419036
c	308	28	2.7	597	13	AQ571789
	309	28	2.7	612	11	BF980214
	310	28	2.7	618	13	B95496
c	311	28	2.7	618	13	AQ543062
	312	28	2.7	622	13	AQ017432
c	313	28	2.7	635	13	AQ425110
	314	28	2.7	644	10	AI207582
	315	28	2.7	650	13	AQ415925
	316	28	2.7	679	13	AQ321467
	317	28	2.7	702	13	AQ478621
	318	28	2.7	724	10	AL048376
	319	28	2.7	736	11	BG482301
	320	28	2.7	755	11	BG481747
	321	28	2.7	759	13	AQ538545
c	322	28	2.7	797	11	BG187463
	323	28	2.7	804	11	BG428355
c	324	28	2.7	805	10	AU120197
	325	28	2.7	828	11	BG432126
	326	28	2.7	855	10	AL157580
	327	28	2.7	861	10	AI525789
	328	28	2.7	879	13	AQ741630
	329	28	2.7	885	11	BF575954
c	330	28	2.7	895	11	BI091119
	331	28	2.7	911	11	BE910447
	332	28	2.7	934	13	AQ750230
c	333	28	2.7	946	10	BE621403
	334	28	2.7	959	11	BG340032
	335	28	2.7	1048	11	BG745833
	336	28	2.7	1129	11	BF340597
c	337	28	2.7	1606	13	AF101950
	338	28	2.7	2225	12	AF130097
	339	27	2.6	105	11	BG951123
	340	27	2.6	105	11	BI041679
c	341	27	2.6	121	11	BF897203
	342	27	2.6	134	11	BF880108
	343	27	2.6	142	10	BE180038
	344	27	2.6	149	11	BG999334
c	345	27	2.6	149	13	AQ533020
	346	27	2.6	153	11	BG957689
c	347	27	2.6	160	10	AW857779
	348	27	2.6	161	11	BG949998
	349	27	2.6	162	11	BF878268
	350	27	2.6	165	10	AW857759
	351	27	2.6	167	10	AA658018
c	352	27	2.6	177	10	AW803364
	353	27	2.6	180	11	BI038819
	354	27	2.6	183	11	BF909161
c	355	27	2.6	190	10	AI349649
	356	27	2.6	199	13	AQ423011
	357	27	2.6	200	10	AA157711
	358	27	2.6	200	10	BE158821
c	359	27	2.6	200	11	BI029295
	360	27	2.6	201	11	BI029152
	361	27	2.6	201	13	B42128
	362	27	2.6	208	11	BF853735
	363	27	2.6	211	10	AA605074
	364	27	2.6	211	13	B01997
c	365	27	2.6	215	13	B55313
	366	27	2.6	219	10	AA229322
	367	27	2.6	220	10	BE158395
	368	27	2.6	222	13	AQ262194
c	369	27	2.6	223	10	BE158833
	370	27	2.6	227	13	AQ626335
	371	27	2.6	228	13	AQ553065
	372	27	2.6	230	13	AQ040316
	373	27	2.6	233	10	BE261473
	374	27	2.6	234	10	AA569786
	375	27	2.6	235	11	T51084
	376	27	2.6	239	11	T36297
	377	27	2.6	240	13	B37604

AQ264917	CITBI-E1-
BI010689	MR2-EN009
AQ419036	RPCI-11-1
AQ571789	HS_2094_A
BF980214	602288048
B95496	CIT-HSP-216
AQ543062	RPCI-11-3
AQ017432	CIT-HSP-2
AQ425110	CITBI-E1-
AI207582	HA2946 Hu
AQ415925	RPCI-11-2
AQ321467	RPCI11-92
AQ478621	RPCI-11-2
AL048376	DKFZp586E
BG482301	602526637
BG481747	602528105
AQ538545	RPCI-11-3
BG187463	RST6456 A
BG428355	602495340
AU120197	AU120197
BG432126	602496201
AL157580	DKFZp761H
AI525789	PT1.3.05
AQ741630	HS_5572_B
BF575954	602132930
BI091119	602854730
BE910447	601503752
AQ750230	HS_5576_A
BE621403	601493826
BG340032	602438190
BG745833	602723980
BF340597	602037348
AF101950	AF101950
AF130097	Homo sapi
BG951123	CMO-CT069
BI041679	IL3-NT029
BF897203	IL2-MT017
BF880108	IL3-ET011
BE180038	IL3-HT061
BG999334	PM0-HN007
BG957689	PM0-CT080
AW857779	RC2-CT032
BG949998	PM2-BN013
BF878268	CM3-ET009
AA658018	nu15c02.s
AW803364	IL2-UM007
BI038819	CM3-NT028
BF909161	PM3-UT005
AI349649	tb58a03.x
AQ423011	CITBI-E1-
AA157711	zo46e03.s
BE158821	IL2-HT039
BI029295	IL5-MT026
BI029152	IL5-MT026
B42128	HS-1055-A2-
BF853735	MR2-EN009
AA605074	no74d02.s
B01997	CSRL-145D11
B55313	CIT-HSP-386
AA229322	nc46c05.s
BE158395	CM0-HT038
AQ262194	CITBI-E1-
BE158833	IL2-HT039
AQ626335	CITBI-E1-
AQ553065	RPCI-11-4
AQ040316	CIT-HSP-2
BE261473	601148613
AA569786	nm41d11.s
T51084	yb71f09.r1
T36297	EST99828 Hu
B37604	HS-1044-B1-

C 378 27 BI014567 241 11 BI014567
C 379 27 BI038137 244 11 BI038137
C 380 27 AQ275929 244 11 AQ275929
C 381 27 BI038809 245 11 BI038809
C 382 27 AW170716 247 10 AW170716
C 383 27 AA730916 248 10 AA730916
C 384 27 BF772982 249 11 BF772982
C 385 27 AA482556 250 10 AA482556
C 386 27 BE464922 253 10 BE464922
C 387 27 AI567372 254 10 AI567372
C 388 27 AW183015 255 10 AW183015
C 389 27 AI367756 256 10 AI367756
C 390 27 AI267161 257 10 AI267161
C 391 27 B87164 257 13 B87164
C 392 27 AQ626001 257 13 AQ626001
C 393 27 AA782405 258 10 AA782405
C 394 27 AI540275 259 10 AI540275
C 395 27 AI371864 261 10 AI371864
C 396 27 AQ108193 262 13 AQ108193
C 397 27 AA458575 263 10 AA458575
C 398 27 AA229255 264 10 AA229255
C 399 27 AL134042 265 10 AL134042
C 400 27 AA658320 266 10 AA658320
C 401 27 BG980642 266 11 BG980642
C 402 27 AW270693 267 10 AW270693
C 403 27 AI758866 268 10 AI758866
C 404 27 AI332831 269 10 AI332831
C 405 27 AA113129 269 10 AA113129
C 406 27 AI719747 269 10 AI719747
C 407 27 R02101 269 11 R02101
C 408 27 T26442 270 11 T26442
C 409 27 AI269212 271 10 AI269212
C 410 27 AI761458 271 10 AI761458
C 411 27 BE158836 271 10 BE158836
C 412 27 N64587 272 11 N64587
C 413 27 AI473842 273 10 AI473842
C 414 27 AV739829 273 10 AV739829
C 415 27 H95865 277 11 H95865
C 416 27 BI010361 278 11 BI010361
C 417 27 M78108 278 11 M78108
C 418 27 AA984754 281 10 AA984754
C 419 27 AQ593014 283 13 AQ593014
C 420 27 AQ017472 285 13 AQ017472
C 421 27 AQ549112 285 13 AQ549112
C 422 27 AL042630 286 10 AL042630
C 423 27 AA984829 287 10 AA984829
C 424 27 T39156 288 11 T39156
C 425 27 AA487892 291 10 AA487892
C 426 27 AQ086725 292 13 AQ086725
C 427 27 B85662 293 13 B85662
C 428 27 BI036077 294 11 BI036077
C 429 27 AA856945 296 10 AA856945
C 430 27 AA883910 296 10 AA883910
C 431 27 AA244303 297 10 AA244303
C 432 27 N70847 297 11 N70847
C 433 27 AQ605614 297 13 AQ605614
C 434 27 BI016488 298 11 BI016488
C 435 27 AZ518668 298 13 AZ518668
C 436 27 AA297207 299 10 AA297207
C 437 27 F32823 301 11 F32823
C 438 27 AI697466 303 10 AI697466
C 439 27 AW731919 304 10 AW731919
C 440 27 AA662452 305 10 AA662452
C 441 27 T52822 308 11 T52822
C 442 27 H24040 309 11 H24040
C 443 27 AA158983 311 10 AA158983
C 444 27 AA584810 311 10 AA584810
C 445 27 AA342457 312 10 AA342457
C 446 27 AA516028 312 10 AA516028
C 447 27 AW834482 313 10 AW834482
C 448 27 AA565555 313 10 AA565555
C 449 27 AA218850 314 10 AA218850
C 450 27 BI044844 314 11 BI044844

BI014567 CM3-ET009
BI038137 CM3-NT028
AQ275929 CITBI-EI-
BI038809 CM3-NT028
AW170716 xi97a08.s
AA730916 nw50e01.s
BF772982 CM3-IT004
AA482556 zt34c06.s
BE464922 hs88b06.x
AI567372 tp67b09.x
AW183015 xp99g07.x
AI367756 qt55g08.x
AI267161 aq49f02.x
B87164 RPCI11-27P2
AQ626001 CITBI-EI-
AA782405 ai31h08.s
AI540275 th46h05.x
AI371864 ta34h06.x
AQ106193 HS_3055_A
AA458575 aa83h02.s
AA229255 nc47e04.s
AL134042 DKFZp547B
AA658320 nu17e05.s
BG980642 CM4-CN008
AW270693 xq56g12.x
AI758866 ty94a06.x
AI332831 qp96f09.x
AA113129 zm29d07.s
AI719747 as92d09.x
R02101 ye83c07.s1
T26442 AB211H9F In
AI269212 qj68g04.x
AI761458 wg65h07.x
BE158836 IL2-HT039
N64587 yz51h09.s1
AI473842 tal7f01.x
AV739829 AV739829
H95865 Yw48c10.s1
BI010361 MR2-EN009
M78108 EST01698 Su
AA984754 am90f04.s
AQ593014 HS_5430_B
AQ017472 CIT-HSP-2
AQ549112 RPCI-11-4
AL042630 DKFZp434M
AA984829 am93f09.s
T39156 ya01c10.r2
AA487892 abl2b07.s
AQ086725 HS_2188_A
B85662 RPCI11-20H1
BI036077 IL2-NT020
AA856945 oe03c08.s
AA883910 aj13a10.s
AA244303 nc14b01.r
N70847 za88a12.s1
AQ605614 HS_5386_A
BI016488 PM0-ET025
AZ518668 RPCI-11-2
AA297207 EST112750
F32823 HSPD25866 H
AI697466 tq08g07.x
AW731919 ba03a11.x
AA662452 nu91f04.s
T52822 ya81a11.s1
H24040 Ym54a07.s1
AA158983 zo57a05.s
AA584810 no10d08.s
AA342457 EST48183
AA516028 ng65a11.s
AW834482 MR2-TT001
AA565555 nk42e06.s
AA218850 zq15b09.s
BI044844 PM0-OT020

C 451 27
C 452 27
C 453 27
C 454 27
C 455 27
C 456 27
C 457 27
C 458 27
C 459 27
C 460 27
C 461 27
C 462 27
C 463 27
C 464 27
C 465 27
C 466 27
C 467 27
C 468 27
C 469 27
C 470 27
C 471 27
C 472 27
C 473 27
C 474 27
C 475 27
C 476 27
C 477 27
C 478 27
C 479 27
C 480 27
C 481 27
C 482 27
C 483 27
C 484 27
C 485 27
C 486 27
C 487 27
C 488 27
C 489 27
C 490 27
C 491 27
C 492 27
C 493 27
C 494 27
C 495 27
C 496 27
C 497 27
C 498 27
C 499 27
C 500 27
C 501 27
C 502 27
C 503 27
C 504 27
C 505 27
C 506 27
C 507 27
C 508 27
C 509 27
C 510 27
C 511 27
C 512 27
C 513 27
C 514 27
C 515 27
C 516 27
C 517 27
C 518 27
C 519 27
C 520 27
C 521 27
C 522 27
C 523 27

N85485 J3701F Huma
AW772696 hn05g04.x
AA663420 ac51c07.s
AA446804 zw80g04.r
BG006858 IL5-GN023
AQ392145 CITBI-EI-
AQ378288 RPCI11-16
AQ388603 RPCI11-15
AQ314621 RPCI11-10
B91498 CIT-HSP-217
AA985332 am79g05.s
AA632640 np87b03.s
AI284467 qu82a03.x
BF886839 PM3-TN018
AA830070 oc44a10.s
BG012742 IL5-GN023
F01304 HSC6E012 S
BF742465 RC1-BT025
AQ081573 RPCI11-54
AA748140 nx79h10.s
AA604632 no84c10.s
BF899031 PM2-MT020
AA196204 zp92c04.s
AQ043469 CIT-HSP-2
BG958931 PM2-CT026
BF873595 IL5-ET011
AQ603977 HS_2136_A
AI475147 tl79b09.x
F08649 HSC1DE111 n
AW179336 PM4-ST012
AA747957 nx68d12.s
AI246884 qx51e12.x
BE069060 QV3-BT037
AQ239144 RPCI11-75
BE068949 RC5-BT037
AW812826 RC3-ST018
AQ047409 CIT-HSP-2
BI031199 IL5-MT020
W21392 zb49a09.r1
AA987606 or93a11.s
AA653041 ns71e11.s
AI734052 nc09f07.y
AW069769 cr50c02.x
AA229034 nc50d10.s
AA573631 nf43b05.s
BF922719 QV4-NT025
AJ346511 SHR-00001
AA533855 nj94f04.s
AV710483 AV710483
BG942919 ax31c04.x
H92152 ys84e09.r1
BF919050 CM0-NT013
AQ281592 RPCI11-88
AA837716 oe06d05.s
BE562811 601336517
BG170207 602321718
AW793601 MR1-UM000
AA678632 ah07a02.s
AI201706 q144c10.x
AQ588377 CITBI-EI-
AA329122 EST32968
AI354409 qu17a02.x
AA515926 ng64g01.s
BG944366 ax49f07.x
BG980850 CM4-CN008
AQ208401 HS_3230_A
AQ263512 CITBI-EI-
AA502098 ng58a03.s
AA506838 nm62f07.s
AW166641 xg54g09.x
T69896 yc18h04.r1
AA400436 zu62b12.s
BF849841 PM4-EN006

670 AL044793 27 2.6 492 10 AL044793
671 27 2.6 494 10 AW471092
672 27 2.6 495 10 AI500640
c 673 27 2.6 498 13 AQ422506
674 27 2.6 499 11 BG777252
c 675 27 2.6 499 11 BF919931
c 676 27 2.6 499 13 B81834
677 27 2.6 500 10 AA496808
678 27 2.6 500 13 AQ620271
679 27 2.6 501 13 AQ226133
680 27 2.6 501 13 AQ489571
c 681 27 2.6 504 10 AL597533
c 682 27 2.6 504 10 AV739938
c 683 27 2.6 504 10 AA626478
c 684 27 2.6 504 13 AQ429961
c 685 27 2.6 505 10 AV739810
c 686 27 2.6 506 10 AA492114
687 27 2.6 507 10 AA400544
c 688 27 2.6 509 13 AQ211741
689 27 2.6 510 10 AA670456
690 27 2.6 514 10 AA482950
c 691 27 2.6 514 13 AQ244398
c 692 27 2.6 515 13 AQ526160
c 693 27 2.6 516 13 AQ512553
c 694 27 2.6 518 13 AQ562887
c 695 27 2.6 519 10 AI732085
c 696 27 2.6 519 13 AQ512745
c 697 27 2.6 520 11 BF923817
c 698 27 2.6 522 10 BE178481
c 699 27 2.6 523 10 BE065970
c 700 27 2.6 524 10 AW167073
c 701 27 2.6 524 11 BF737749
c 702 27 2.6 525 11 BI039053
c 703 27 2.6 526 11 BI039054
c 704 27 2.6 526 11 BG009072
c 705 27 2.6 529 10 AU144449
c 706 27 2.6 529 10 AV720925
c 707 27 2.6 531 13 AQ536676
c 708 27 2.6 532 13 AQ691213
c 709 27 2.6 533 11 BG056397
c 710 27 2.6 533 13 AQ708350
c 711 27 2.6 534 13 AQ080837
c 712 27 2.6 535 10 AL042731
c 713 27 2.6 536 10 AU146724
c 714 27 2.6 536 11 BI021436
c 715 27 2.6 536 13 AQ480181
c 716 27 2.6 537 13 AQ332193
c 717 27 2.6 538 11 BI039969
c 718 27 2.6 538 11 BF448499
c 719 27 2.6 541 10 AW975367
c 720 27 2.6 542 11 H23506
c 721 27 2.6 542 13 AZ254674
c 722 27 2.6 543 11 BF826816
c 723 27 2.6 545 10 AW939214
c 724 27 2.6 549 13 AQ570204
c 725 27 2.6 550 13 AQ721483
c 726 27 2.6 550 13 AQ352086
c 727 27 2.6 551 13 AQ000610
c 728 27 2.6 551 13 AQ020745
c 729 27 2.6 552 10 BE740827
c 730 27 2.6 552 11 BG983313
c 731 27 2.6 552 11 N78063
c 732 27 2.6 552 13 AQ678716
c 733 27 2.6 553 10 BE704103
c 734 27 2.6 553 13 AQ334955
c 735 27 2.6 555 13 AQ533504
c 736 27 2.6 556 11 BF675599
c 737 27 2.6 557 10 BE143857
c 738 27 2.6 557 13 AQ354690
c 739 27 2.6 558 11 BF954499
c 740 27 2.6 559 10 BE299730
c 741 27 2.6 559 13 AQ441794
c 742 27 2.6 561 10 AW603818

AL044793 DKFZp434N
AW471092 xu06h04.x
AI500640 tn93e02.x
AQ422506 RPCI-11-1
BG777252 602664483
BF919931 QV0-NT015
B81834 RPCI11-17I6
AA496808 ae33h09.s
AQ620271 HS_5182_B
AQ226133 HS_2015_B
AQ489571 RPCI-11-2
AL597533 DKFZp313G
AV739938 AV739938
AA626478 ab53h12.r
AQ429961 HS_5077_A
AV739810 AV739810
AA492114 ng17d03.s
AA400544 zu62b12.r
AQ211741 HS_3239_A
AA670456 ae62h06.s
AA482950 nf49e09.s
AQ244398 HS_2216_B
AQ526160 HS_5312_B
AQ512553 HS_5119_B
AQ562887 HS_2069_B
AI732085 nc11f02.x
AQ512745 HS_5142_A
BF923817 MR4-NT014
BE178481 RC3-HT060
BE065970 RC3-BT031
AW167073 xg70f01.x
BF737749 PM0-KT000
BI039053 MR4-NT014
BI039054 MR4-NT014
BG009072 RC1-GN019
AU144449 AU144449
AV720925 AV720925
AQ536676 RPCI-11-3
AQ691213 HS_5338_A
BG056397 naf68b04.
AQ708350 HS_5569_A
AQ080837 HS_3171_A
AL042731 DKFZp434B
AU146724 AU146724
BI021436 IL3-NT026
AQ480181 RPCI-11-2
AQ332193 HS_5003_A
BI039969 MR4-NT014
BF448499 7n71e09.x
AW975367 EST387475
H23506 ym57h11.s1
AZ254674 UP_561-7J
BF826816 MR2-HN003
AW939214 QV1-DT007
AQ570204 HS_5356_A
AQ721483 HS_5552_B
AQ352086 CITBI-E1-
AQ000610 CIT-HSP-2
AQ020745 CIT-HSP-2
BE740827 601593034
BG983313 IL5-CN006
N78063 yv72a01.r1
AQ678716 HS_2094_B
BE704103 MR3-NN112
AQ334955 HS_5012_B
AQ533504 RPCI-11-3
BF675599 602083488
BE143857 MR0-HT016
AQ354690 CITBI-E1-
BF954499 MR4-NN020
BE299730 600944380
AQ441794 HS_5106_B
AW603818 CM0-CN004

c 743 27
c 744 27
745 27
c 746 27
747 27
748 27
749 27
750 27
751 27
c 752 27
753 27
754 27
755 27
756 27
757 27
758 27
759 27
760 27
761 27
762 27
763 27
c 764 27
765 27
c 766 27
c 767 27
c 768 27
c 769 27
c 770 27
c 771 27
c 772 27
773 27
c 774 27
c 775 27
c 776 27
777 27
c 778 27
c 779 27
c 780 27
c 781 27
c 782 27
c 783 27
c 784 27
785 27
786 27
787 27
c 788 27
c 789 27
790 27
c 791 27
c 792 27
c 793 27
c 794 27
c 795 27
c 796 27
c 797 27
798 27
799 27
c 800 27
c 801 27
c 802 27
c 803 27
804 27
c 805 27
806 27
807 27
c 808 27
c 809 27
c 810 27
c 811 27
c 812 27
c 813 27
c 814 27
815 27

AQ562764 HS_2079_B
AW889786 RC1-NT003
AU147926 AU147926
AQ568574 HS_2081_A
AU147352 AU147352
N25784 yx84e04.s1
AQ791303 HS_4555_B
AW409893 fh03a03.y
AQ547583 RPCI-11-4
AQ346926 RPCI11-11
BE884181 601505936
AW965632 EST377705
AW975952 EST388061
AW965633 EST377706
AL046677 DKFZp434B
BI000101 MR3-HN006
BF941382 7d98b09.x
AV718803 AV718803
B53056 CIT-HSP-200
AA428512 zw47a04.r
AQ636023 RPCI-11-4
N58473 yv72a01.s1
AL577347 AL577347
AQ321791 RPCI11-93
BE160626 PM1-HT042
AV723671 AV723671
BE143624 MR0-HT016
AG019176 Homo sapi
BF905405 IL3-MT026
AQ383938 RPCI11-13
AA155784 zo49e12.s
W81268 zd85g07.r1
BG570417 602590861
AW603838 CM0-CN004
AA133332 zn92e02.s
BG506930 601861555
AV764087 AV764087
AW853217 RC1-CT024
AQ666530 HS_2096_A
AQ939447 NR1-151C
AW853213 RC1-CT024
AV702975 AV702975
AQ040460 CIT-HSP-2
AW961953 EST374026
AQ279584 CITBI-E1-
AQ312956 RPCI11-10
AG025882 Homo sapi
BE147973 RC3-HT023
AQ267293 RPCI11-68
AQ587626 CITBI-E1-
AG025922 Homo sapi
AI133514 HA2127 Hu
BG392409 602410579
BG542392 602571759
AW962959 EST375032
AQ039461 CIT-HSP-2
AQ314145 RPCI11-10
AV758902 AV758902
AG019084 Homo sapi
AQ470989 CITBI-E1-
BG202532 RST21891
AW969652 EST381729
AV763950 AV763950
BE273856 601104362
AQ528747 RPCI-11-3
AQ037792 CIT-HSP-2
AQ630943 RPCI-11-4
AV760878 AV760878
AV704876 AV704876
BE545198 601074970
AQ351773 CITBI-E1-
BG536089 602564367
AQ543328 RPCI-11-3

13 AQ562764
561 13 AQ562764
562 10 AW889786
563 10 AU147926
565 13 AQ568574
566 10 AU147352
566 11 N25784
568 13 AQ791303
571 10 AW409893
572 13 AQ547583
573 13 AQ346926
574 11 BE884181
580 10 AW965632
581 10 AW975952
582 10 AW965633
585 10 AL046677
585 11 BI000101
585 11 BF941382
586 10 AV718803
587 13 B53056
588 10 AA428512
588 13 AQ636023
591 11 N58473
592 10 AL577347
592 13 AQ321791
593 10 BE160626
594 10 AV723671
594 10 BE143624
598 13 AG019176
599 11 BF905405
601 13 AQ383938
603 10 AA155784
605 11 W81268
606 11 BG570417
610 10 AW603838
612 10 AA133332
613 11 BG506930
615 10 AV764087
615 10 AW853217
615 13 AQ666530
616 13 AQ939447
620 10 AW853213
623 10 AV702975
628 13 AQ040460
634 10 AW961953
635 13 AQ279584
636 13 AQ312956
637 13 AG025882
640 10 BE147973
644 13 AQ267293
646 13 AQ587626
648 13 AG025922
651 10 AI133514
651 11 BG392409
653 11 BG542392
657 10 AW962959
657 13 AQ039461
657 13 AQ314145
658 10 AV758902
659 13 AG019084
659 13 AQ470989
662 11 BG202532
663 10 AW969652
672 10 AV763950
672 10 BE273856
673 13 AQ528747
676 13 AQ037792
676 13 AQ630943
681 10 AV760878
686 10 AV704876
689 10 BE545198
690 13 AQ51773
692 11 BG536089
692 13 AQ543328

c 816	27	2.6	696	10	AV718208	AV718208	AV718208	AV718208	889	27	2.6	1071	10	AL569845	AL569845
c 817	27	2.6	696	13	AG016448	AG016448	Homo sapi	AG016448	c 890	27	2.6	1079	11	BG698721	BG698721
c 818	27	2.6	697	10	AV762945	AV762945	AV762945	AV762945	c 891	27	2.6	1091	11	BF681181	BF681181
c 819	27	2.6	697	13	AG016444	AG016444	Homo sapi	AG016444	c 892	27	2.6	1129	11	BG775312	BG775312
c 820	27	2.6	698	10	AG0117514	AG0117514	Homo sapi	AG0117514	c 893	27	2.6	1172	12	AF305824	AF305824
c 821	27	2.6	699	10	BE169990	BE169990	CM0-HT053	BE169990	c 894	27	2.6	1187	11	BG251369	BG251369
c 822	27	2.6	699	13	AO017056	AO017056	CIT-HSP-2	AO017056	c 895	27	2.6	1243	10	AV763002	AV763002
c 823	27	2.6	699	13	AG013234	AG013234	Homo sapi	AG013234	c 896	27	2.6	1247	10	AV761200	AV761200
c 824	27	2.6	704	13	AG016449	AG016449	Homo sapi	AG016449	c 897	27	2.6	1335	10	AV761381	AV761381
c 825	27	2.6	704	13	AO543793	AO543793	RPCI-11-3	AO543793	c 898	27	2.6	1474	10	AV762221	AV762221
c 826	27	2.6	705	10	AV707907	AV707907	AV707907	AV707907	c 899	27	2.6	1639	12	AF113684	AF113684
c 827	27	2.6	709	13	AG013216	AG013216	Homo sapi	AG013216	c 900	27	2.6	1898	12	AF119861	AF119861
c 828	27	2.6	711	11	BF666079	BF666079	602118862	BF666079	c 901	26	2.5	168	10	AA828137	AA828137
c 829	27	2.6	714	10	AV713462	AV713462	AV713462	AV713462	c 902	26	2.5	169	13	B55051	B55051
c 830	27	2.6	716	13	AG013215	AG013215	Homo sapi	AG013215	c 903	26	2.5	182	13	AO479538	AO479538
c 831	27	2.6	718	10	AW963636	AW963636	EST375709	AW963636	c 904	26	2.5	194	10	AA658182	AA658182
c 832	27	2.6	720	11	BF968561	BF968561	602271163	BF968561	c 905	26	2.5	197	11	T51743	T51743
c 833	27	2.6	723	11	BF902627	BF902627	602549457	BF902627	c 906	26	2.5	201	11	BF886789	BF886789
c 834	27	2.6	729	10	AV762685	AV762685	AV762685	AV762685	c 907	26	2.5	205	10	AW081911	AW081911
c 835	27	2.6	730	10	AV760533	AV760533	AV760533	AV760533	c 908	26	2.5	214	11	BF991590	BF991590
c 836	27	2.6	730	11	BG529467	BG529467	602557846	BG529467	c 909	26	2.5	216	11	R24500	R24500
c 837	27	2.6	730	11	BF666566	BF666566	602121426	BF666566	c 910	26	2.5	218	11	T12423	T12423
c 838	27	2.6	732	11	BG289613	BG289613	602381582	BG289613	c 911	26	2.5	218	13	B71710	B71710
c 839	27	2.6	733	11	BG675576	BG675576	602622177	BG675576	c 912	26	2.5	221	11	N70241	N70241
c 840	27	2.6	734	11	BG534528	BG534528	602553772	BG534528	c 913	26	2.5	226	10	AA344084	AA344084
c 841	27	2.6	736	11	BG179169	BG179169	602330045	BG179169	c 914	26	2.5	226	10	BE092802	BE092802
c 842	27	2.6	737	10	AL601138	AL601138	DKF2p313P	AL601138	c 915	26	2.5	233	10	AW905647	AW905647
c 843	27	2.6	742	10	AV762133	AV762133	AV762133	AV762133	c 916	26	2.5	233	11	BF813493	BF813493
c 844	27	2.6	742	13	AG011957	AG011957	Homo sapi	AG011957	c 917	26	2.5	234	11	F34457	F34457
c 845	27	2.6	745	13	AQ895263	AQ895263	HS_3148_B	AQ895263	c 918	26	2.5	262	10	AI678542	AI678542
c 846	27	2.6	751	11	BI084691	BI084691	602869885	BI084691	c 919	26	2.5	263	10	AW937369	AW937369
c 847	27	2.6	752	10	AV758960	AV758960	AV758960	AV758960	c 920	26	2.5	266	10	AA827295	AA827295
c 848	27	2.6	753	10	BE299494	BE299494	600944773	BE299494	c 921	26	2.5	266	13	AO677178	AO677178
c 849	27	2.6	753	11	BF671285	BF671285	602151145	BF671285	c 922	26	2.5	268	10	AA327534	AA327534
c 850	27	2.6	755	13	AG011956	AG011956	Homo sapi	AG011956	c 923	26	2.5	270	10	AA303165	AA303165
c 851	27	2.6	761	10	AV761908	AV761908	AV761908	AV761908	c 924	26	2.5	272	10	AI339394	AI339394
c 852	27	2.6	761	11	BG681910	BG681910	602629849	BG681910	c 925	26	2.5	273	11	BF914788	BF914788
c 853	27	2.6	765	11	BF036900	BF036900	601459106	BF036900	c 926	26	2.5	275	10	AA635067	AA635067
c 854	27	2.6	765	13	AQ896662	AQ896662	HS_3117_B	AQ896662	c 927	26	2.5	295	10	AA376615	AA376615
c 855	27	2.6	769	10	AL521028	AL521028	AL521028	AL521028	c 928	26	2.5	302	10	AA838096	AA838096
c 856	27	2.6	770	10	AL582702	AL582702	AL582702	AL582702	c 929	26	2.5	302	10	AW934751	AW934751
c 857	27	2.6	782	11	BG434786	BG434786	602507301	BG434786	c 930	26	2.5	305	11	BG997683	BG997683
c 858	27	2.6	794	11	BG740010	BG740010	602630973	BG740010	c 931	26	2.5	312	10	AW805169	AW805169
c 859	27	2.6	829	11	BG911559	BG911559	602808862	BG911559	c 932	26	2.5	314	10	AW085785	AW085785
c 860	27	2.6	829	11	BG113515	BG113515	602283973	BG113515	c 933	26	2.5	316	13	B78760	B78760
c 861	27	2.6	837	11	BF692882	BF692882	60280310	BF692882	c 934	26	2.5	319	10	AW837332	AW837332
c 862	27	2.6	842	13	BO2326	BO2326	cSRL-151E8-	BO2326	c 935	26	2.5	325	10	BE139033	BE139033
c 863	27	2.6	845	10	AI557106	AI557106	PT2_1_13-	AI557106	c 936	26	2.5	326	11	BI062186	BI062186
c 864	27	2.6	851	11	BG283407	BG283407	602407167	BG283407	c 937	26	2.5	326	13	AQ090479	AQ090479
c 865	27	2.6	858	11	BF214379	BF214379	601845468	BF214379	c 938	26	2.5	327	10	AA405517	AA405517
c 866	27	2.6	859	11	BG615899	BG615899	602643318	BG615899	c 939	26	2.5	344	10	AW892026	AW892026
c 867	27	2.6	891	11	BE972613	BE972613	601652510	BE972613	c 940	26	2.5	346	10	AA729166	AA729166
c 868	27	2.6	906	10	BE784083	BE784083	601472577	BE784083	c 941	26	2.5	347	10	AA847357	AA847357
c 869	27	2.6	914	11	BG912185	BG912185	602812873	BG912185	c 942	26	2.5	352	11	R89367	R89367
c 870	27	2.6	917	11	BE85138	BE85138	601510892	BE85138	c 943	26	2.5	354	10	AW817527	AW817527
c 871	27	2.6	924	11	BG388300	BG388300	602413307	BG388300	c 944	26	2.5	355	10	AI420666	AI420666
c 872	27	2.6	924	11	BG697538	BG697538	602660758	BG697538	c 945	26	2.5	356	13	BI5838	BI5838
c 873	27	2.6	935	11	BG432843	BG432843	602496051	BG432843	c 946	26	2.5	357	13	AQ665282	AQ665282
c 874	27	2.6	935	11	BG677896	BG677896	602625591	BG677896	c 947	26	2.5	358	11	BF843520	BF843520
c 875	27	2.6	943	10	AL527924	AL527924	AL527924	AL527924	c 948	26	2.5	360	10	AA468990	AA468990
c 876	27	2.6	946	10	AL570329	AL570329	AL570329	AL570329	c 949	26	2.5	364	10	AA468990	AA468990
c 877	27	2.6	956	11	BG110329	BG110329	60279216	BG110329	c 950	26	2.5	364	10	AW861716	AW861716
c 878	27	2.6	969	11	BG615087	BG615087	601652510	BG615087	c 951	26	2.5	366	10	AA501867	AA501867
c 879	27	2.6	970	10	BE788445	BE788445	601475624	BE788445	c 952	26	2.5	366	10	AV762966	AV762966
c 880	27	2.6	976	10	AL553425	AL553425	AL553425	AL553425	c 953	26	2.5	368	11	H82863	H82863
c 881	27	2.6	978	11	BF027503	BF027503	601671778	BF027503	c 954	26	2.5	369	10	AA618172	AA618172
c 882	27	2.6	989	11	BG116144	BG116144	602318312	BG116144	c 955	26	2.5	369	11	AA618172	AA618172
c 883	27	2.6	990	10	AL525925	AL525925	AL525925	AL525925	c 956	26	2.5	372	10	AW937790	AW937790
c 884	27	2.6	1005	10	AL566411	AL566411	AL566411	AL566411	c 957	26	2.5	373	13	AQ278367	AQ278367
c 885	27	2.6	1005	11	BG576422	BG576422	602595661	BG576422	c 958	26	2.5	375	10	AA507474	AA507474
c 886	27	2.6	1024	11	BG706247	BG706247	602669544	BG706247	c 959	26	2.5	376	10	AW827426	AW827426
c 887	27	2.6	1027	11	BG622059	BG622059	602646665	BG622059	c 960	26	2.5	377	13	AQ631460	AQ631460
c 888	27	2.6	1036	10	BE299695	BE299695	600944380	BE299695	c 961	26	2.5	383	11	BG152452	BG152452
	27	2.6	1036	10	BE299695	BE299695	600944380	BE299695		26	2.5	394	11	BI034667	BI034667

c 962 26 2.5 394 11 BF963551 PM1-NN120
963 26 2.5 399 10 AA679235 ac57f11.s
964 26 2.5 399 10 AA962011 oq92f05.s
965 26 2.5 399 10 AI888365 w21c10.x
c 966 26 2.5 401 10 AV738723 AV738723
967 26 2.5 403 10 AA973261 cl89h06.s
c 968 26 2.5 403 10 AA251241 zs03a04.s
c 969 26 2.5 404 13 AQ009396 RPII11-23
970 26 2.5 406 11 T03750 IB846 Infan
971 26 2.5 406 13 AQ036911 CIT-HSP-2
c 972 26 2.5 408 11 BF872591 IL3-ET011
973 26 2.5 411 11 H53579 yq87c08.sl
974 26 2.5 412 10 AA835938 oc77e07.s
c 975 26 2.5 412 10 AI862153 ta10f12.x
c 976 26 2.5 414 11 BF943816 RC1-NN019
c 977 26 2.5 415 11 T95842 ye42c07.s1
978 26 2.5 418 10 AA081047 zn33e04.s
979 26 2.5 418 10 AA310218 EST18105
980 26 2.5 418 11 BF901094 IL2-MT017
981 26 2.5 419 10 AI520724 ti06g06.x
982 26 2.5 420 13 B99694 CIT-HSP-228
c 983 26 2.5 421 13 AQ060139 CIT-HSP-2
984 26 2.5 423 10 AA113126 AA113126 zm29d04.s
985 26 2.5 425 10 AA436716 zv59d07.s
c 986 26 2.5 425 13 AZ520178 RPII-11-1
987 26 2.5 427 10 AA121862 zn96c09.r
988 26 2.5 427 10 AI922408 wn91g12.x
c 989 26 2.5 427 11 BF954832 PM4-NN120
c 990 26 2.5 428 10 AA632695 np83b02.s
c 991 26 2.5 429 11 BF964095 PM4-NN120
c 992 26 2.5 430 10 AI474110 tm05d09.x
993 26 2.5 430 10 AA598607 ae39f06.s
994 26 2.5 430 11 R27323 yh41e12.s1
c 995 26 2.5 431 10 AI055882 ow30b12.s
c 996 26 2.5 433 13 AQ114206 CIT-HSP-2
997 26 2.5 434 10 AA811150 oa98d05.s
c 998 26 2.5 435 10 AW969798 EST381876
c 999 26 2.5 436 10 AA649183 ns31a11.s
1000 26 2.5 436 10 AA775201 ac79c08.s

ALIGNMENTS

RESULT 1

AA626859 295 bp mRNA EST 15-OCT-1997
zu89f10.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:745195
3' similar to SW:KKIA_HUMAN Q00532 SERINE/THREONINE-PROTEIN KINASE
KKIALRE ; mRNA sequence.
ACCESSION AA626859.1 GI:2539246
VERSION AA626859
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 295)
AUTHORS Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40ml3 fwd. ET from Amersham

FEATURES
source High quality sequence stop: 292.
Location/Qualifiers
1. .295
/organism="Homo sapiens"
/db_xref="GDB:5933055"
/db_xref="taxon:9606"
/clone="IMAGE:745195"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGGAGCGGCCCAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 95 a 69 c 55 g 76 t
ORIGIN

Query Match 19.0%; Score 198; DB 10; Length 295;
Best Local Similarity 100.0%; Pred. No. 1.5e-90;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 455 ttccaggagatgcctacacccgattatgtagctacgagatggtaccgagctcctgaacttc 514
|||||
Db 263 TTCCAGGAGATGCTACACCGATTATGTAGCTACGAGATGCTACCGAGCTCCTGAACCTC 204
|||||
QY 515 ttgtgggagatactcagtcagttcttcagtcagatataatggctattggtgtgttttg 574
|||||
Db 203 TTGTGGGAGATACCTCAGTATGTTCTTCAGTCGATATATGGCTATTTGGTTGTGTTTGTG 144
|||||
QY 575 cagagctcctgacagcagccactgtgacctggcctggaaatcagatgtggaccaactttatc 634
|||||
Db 143 CAGAGCTCCTGACAGGCGCAGCCACTGTGGCCTGGAAATCAGATGTGGACCACTTTATC 84
|||||

QY 635 tgataatcagaacactag 652
|||||
Db 83 TGATAATCAGAACACTAG 66
|||||

RESULT 2
BG390099
LOCUS BG390099 756 bp mRNA EST 12-MAR-2001
DEFINITION 602415940F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4524513 5',
mRNA sequence.
ACCESSION BG390099
VERSION BG390099.1 GI:13283645
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 756)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10428 row: 0 column: 10
High quality sequence stop: 748.
Location/Qualifiers

```
source
1. .756
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4524513"
/clone_lib="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT      205 a   177 c   192 g   182 t
ORIGIN

Query Match      4.5%; Score 47; DB 11; Length 756;
Best Local Similarity 100.0%; Pred. No. 5e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 650 tagtagagacggggttcgccatgttgaccaggctgtctcgaaactc 696
|||||
Db 7 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 53

RESULT 3
BF869052/c
LOCUS
DEFINITION
ACCESSION BF869052
VERSION BF869052.1 GI:12259182
KEYWORDS EST.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 433)
Dias Neto,E., Garcia Correa,R., Verjovsky-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=IL5&t2=IL5-ET0119-
181000-180-f09&t3=2000-10-18&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 396.
Location/Qualifiers
1. .433
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="ET0119"
/dev_stage="Adult"
/notes="Organ: lung_tumor; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
```

```
BASE COUNT      137 a   102 c   111 g   83 t
ORIGIN

Query Match      3.9%; Score 41; DB 11; Length 433;
Best Local Similarity 100.0%; Pred. No. 6.2e-10;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 650 tagtagagacggggttcgccatgttgaccaggctgtctc 690
|||||
Db 124 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTC 84

RESULT 4
AQ230677
LOCUS
DEFINITION
ACCESSION AQ230677
VERSION AQ230677.1 GI:3655906
KEYWORDS GSS.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 328)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
99380589
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 2033 row: B column: 20
Class: BAC ends
High quality sequence stop: 328.
Location/Qualifiers
1. .328
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=2033 Col=20 Row=B"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/notes="Organ: sperm; Vector: pBelobAC11; BAC Clones in
E-Coli DH10B"

BASE COUNT      67 a   94 c   69 g   98 t
ORIGIN

Query Match      3.7%; Score 39; DB 13; Length 328;
Best Local Similarity 100.0%; Pred. No. 6.7e-09;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 650 tagtagagacggggttcgccatgttgaccaggctggtc 688
|||||
Db 232 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTC 270

RESULT 5
BI040145
LOCUS
DEFINITION RC5-NT0263-010201-012-F07 NT0263 Homo sapiens cDNA, mRNA sequence.
ACCESSION BI040145
VERSION BI040145.1 GI:14446771
EST 14-JUN-2001
```


/db_xref="taxon:9606"
/clone="plate=CT 789 Col=17 Row=L"
/clone_lib="CIT Human Genomic Sperm Library C"
/sex="M"
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
E-Coli DH10B"
100 a 123 c 118 g 148 t 5 others
BASE COUNT
ORIGIN

Query Match 3.6%; Score 37; DB 13; Length 494;
Best Local Similarity 100.0%; Pred. No. 7e-08;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 660 ggggttcgccatgttgaccaggctgtctcgaaactc 696
|||||
Db 320 GGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 356

RESULT 8
AG014845/c
LOCUS
DEFINITION
AG014845 Homo sapiens genomic DNA, 21q region, clone: 289H18N11, genomic survey sequence.
AG014845 AG006560
AG014845.1 GI:3650089
KEYWORDS
SOURCE
ORGANISM
Homo sapiens DNA, clone:289H18N11.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 700)
Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
Homo sapiens genomic DNA, chromosome 21q
Published Only in DataBase (1998) In press
2 (bases 1 to 700)
Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
Direct Submission
Submitted (23-SEP-1998) to the DDBJ/EMBL/GenBank databases.
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan
(E-mail:hattori@hg.c.ims.u-tokyo.ac.jp, Tel:0427-78-9732,
Fax:0427-78-9561)
On Feb 6, 1999 this sequence version replaced gi:2992438.
AG006560: Submitted (27-Mar-1998).
FEATURES
source
Location/Qualifiers
1..700
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/clone="289H18N11"
/map="21q"
BASE COUNT 134 a 152 c 156 g 137 t 121 others
ORIGIN

Query Match 3.6%; Score 37; DB 13; Length 700;
Best Local Similarity 100.0%; Pred. No. 6.9e-08;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 660 ggggttcgccatgttgaccaggctgtctcgaaactc 696
|||||
Db 218 GGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 182

RESULT 9
BE150038
LOCUS
DEFINITION
BE150038 QV3-HT0262-041199-014-b05 HT0262 Homo sapiens cDNA, mRNA sequence.
ACCESSION
BE150038
BE150038.1 GI:8612759
KEYWORDS
EST.
human.
human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 207)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=QV3-HT0262-041
199-014-b05&t3=1999-11-04&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 5
High quality sequence stop: 207.
FEATURES
Location/Qualifiers
1..207
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0262"
/dev_stage="Adult"
/note="Organ: head_neck; Vector: puc18; Site.1: SmaI;
Site.2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 45 a 58 c 52 g 50 t 2 others
ORIGIN

Query Match 3.5%; Score 36; DB 10; Length 207;
Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 655 gagacggggttcgccatgttgaccaggctgtctc 690
|||||
Db 67 GAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTC 102

RESULT 10
H95588/c
LOCUS
DEFINITION
H95588 366 bp mRNA EST 25-NOV-1996
Yv16f01.s1 Soares fetal liver spleen lNFLS Homo sapiens cDNA clone
IMAGE:242905 3' similar to contains Alu repetitive element;; mRNA
sequence.
H95588
H95588.1 GI:1108730
KEYWORDS
EST.
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 366)
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlffing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
R., Williamson,A., Wohldmann,P. and Wilson,R.

TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 232
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1759 Std Error: 0.00
Seq primer: ml3 -40 forward.
Location/Qualifiers
1. .366
/organism="Homo sapiens"
/db_xref="GDB:3792038"
/db_xref="taxon:9606"
/clone="IMAGE:242905"
/clone_lib="Soares fetal liver spleen 1NFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5' AACTGGAAGAATTAATAAGATCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 89 a 78 c 107 g 91 t 1 others
ORIGIN

Query Match 3.5%; Score 36; DB 11; Length 366;
Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 661 gggttcgccatgttgaccaggctggctcgaaactc 696
|||||
Db 197 GGGTTTCGCATGTTGACCAAGCTGGTCTCGAACTC 162

RESULT 11
AA678313
LOCUS 408 bp mRNA EST 19-DEC-1997
DEFINITION zill6c01.s1 Soares_fetal_liver_spleen_1NFLS_S1 Homo sapiens cDNA
clone IMAGE:430944 3' similar to contains Alu repetitive element;;
mRNA sequence.
ACCESSION AA678313
VERSION AA678313.1 GI:2658835
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 408)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 355.
Location/Qualifiers
1. .408
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:430944"
/clone_lib="Soares_fetal_liver_spleen_1NFLS_S1"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
This is a subtracted version of the original Soares fetal
liver spleen 1NFLS library. 1st strand cDNA was primed
with a Pac I - oligo(dT) primer [5'
AACTGGAAGAATTAATAAGATCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 85 a 108 c 99 g 116 t
ORIGIN

Query Match 3.5%; Score 36; DB 10; Length 408;
Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 661 gggttcgccatgttgaccaggctggctcgaaactc 696
|||||
Db 177 GGGTTTCGCATGTTGACCAAGCTGGTCTCGAACTC 212

RESULT 12
AQ769170/c
LOCUS 500 bp DNA GSS 28-JUL-1999
DEFINITION HS_3155_B2_B09_T7C CIT Approved Human Genomic Sperm Library D Homo
sapiens genomic clone Plate=3155 Col=18 Row=D, DNA sequence.
ACCESSION AQ769170
VERSION AQ769170.1 GI:5647286
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 500)
AUTHORS Mahairas,G.C., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end Web Server: http://www.htsc.washington.edu
Plate: 3155 row: D column: 18
Seq primer: T7
Class: BAC ends
High quality sequence stop: 500.
Location/Qualifiers
1. .500
/organism="Homo sapiens"
/db_xref="taxon:9606"

FEATURES
source

/clone="plate=3155 Col=18 Row=D"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
E-Coli DH10B"

BASE COUNT 143 a 110 c 131 g 110 t 6 others
ORIGIN

Query Match 3.5%; Score 36; DB 13; Length 500;
Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 661 ggggttcgcatgttgaccaggctggctcgaaactc 696
|||||
Db 330 GGGTTTCGCATGTTGACCAGGCTGGTCTCGAACTC 295

RESULT 13
AQ668860/c
LOCUS
DEFINITION HS_5431_A1_E09_SP6E RPCI-11 Human Male BAC Library Homo sapiens
genomic clone Plate=1007 Col=17 Row=I, DNA sequence.
AQ668860
AQ668860.1 GI:5201606
GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 516)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
99380589
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Resear h Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
plate: 1007 row: I column: 17
Seq primer: SP6
Class: BAC ends
High quality sequence stop: 516.
Location/Qualifiers
1. .516
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="plate=1007 Col=17 Row=I"
/clone_lib="RPCI-11 Human Male BAC Library"
/sex="male"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBACe3.6 vector at EcoRI sites"

BASE COUNT 164 a 107 c 103 g 137 t 5 others
ORIGIN

Query Match 3.5%; Score 36; DB 13; Length 516;
Best Local Similarity 100.0%; Pred. No. 2.3e-07;

Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 654 agagacgggggttcgccatgttgaccaggctgggtct 689
|||||
Db 339 AGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCT 304

RESULT 14
BG698963/c
LOCUS
DEFINITION BG698963 801 bp mRNA EST 07-MAY-2001
602703384F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4800452 5',
mRNA sequence.
ACCESSION BG698963
VERSION BG698963.1 GI:13966783
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 801)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10691 row: h column: 21
High quality sequence stop: 768.
Location/Qualifiers
1. .801
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4800452"
/clone_lib="NCI_CGAP_Skn3"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 233 a 193 c 198 g 177 t
ORIGIN

Query Match 3.5%; Score 36; DB 11; Length 801;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacggggttcgccatgttgaccaggctg 685
|||||
Db 264 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTG 229

RESULT 15
BE149958
LOCUS
DEFINITION BE149958 193 bp mRNA EST 21-JUN-2000
RCO-HT0259-201199-031-e09 HT0259 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE149958
VERSION BE149958.1 GI:8612679
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 193)
Dias Neto,E., Garcia Correa,R., Verjovsky-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,

Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL MEDLINE COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

{http://www.ludwig.org.br/scripts/gethtml2.pl?tl1=&t2=RC0-HT0259-201

199-031-e09&t3=1999-11-20&t4=1}

Seq primer: puc 18 forward

High quality sequence start: 115

High quality sequence stop: 193.

Location/Qualifiers

1. .193

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="HT0259"

/dev_stage="Adult"

/note="Organ: head_neck; Vector: puc18; Site_1: SmaI;

Site_2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the pUC 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Clones are derived from the human BAC library RPCI-11. For BAC

library availability, please contact Pieter de Jong

(pieter@dejong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)

or from Resear h Genetics (info@resgen.com). BAC end Web Server:

http://www.htsc.washington.edu

Plate: 737 row: F column: 14

Seq primer: T7

Class: BAC ends

High quality sequence stop: 501.

Location/Qualifiers

1. .501

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="Plate=737 Col=14 Row=F"

/clone_lib="RPCI-11 Human Male BAC Library"

/sex="male"

/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;

Male blood DNA was isolated from one randomly chosen donor

and partially digested with a combination of EcoRI and

EcoRI Methylase. Size selected DNA was cloned into the

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

pBACe3.6 vector at EcoRI sites"

(Clontech); Site_1: SfiI (ggcgcctcgcc); Site_2: SfiI (ggccattatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGCGGCGCCGACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.55 kb (range 1.0-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

BASE COUNT 161 a 163 c 156 g 240 t
ORIGIN

Query Match 3.4%; Score 35; DB 11; Length 720;
Best Local Similarity 100.0%; Pred. No. 7.3e-07;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 662 ggttcgccatgttgaccaggctgggtctcgaactc 696
|||||
Db 95 GGTTCGCCATGTTGACCAGGCTGGTCTCGAAGTC 129

RESULT 18

AQ3211133 778 bp DNA GSS 06-MAY-1999
LOCUS
DEFINITION RPCI11-111E3.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-111E3, DNA sequence.

ACCESSION AQ3211133
VERSION AQ3211133.1 GI:4053867
KEYWORDS GSS.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 778)
Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C. Use of human BAC End Sequences for Sequence-Ready Map Building Unpublished (1998)

TITLE

Other_GSSs: RPCI11-111E3.TV

JOURNAL

Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208

Email: hbe@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.

FEATURES

source

Location/Qualifiers
1..778
/organism="Homo sapiens"
/db_xref="GDB:7542338"
/db_xref="taxon:9606"
/clone="RPCI-11-111E3"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; RPCI11 Human Male BAC Library"

BASE COUNT

ORIGIN

176 a 185 c 146 g 271 t

Query Match 3.4%; Score 35; DB 13; Length 778;
Best Local Similarity 100.0%; Pred. No. 7.3e-07;

Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 650 tagtagacgggtttcgccatgttgaccaggct 684
|||||
Db 525 TAGTAGACGGGTTTCGCCATGTTGACCAGGCT 559

RESULT 19

BF878219 369 bp mRNA EST 17-JAN-2001
LOCUS
DEFINITION CM2-ET0127-181100-523-fl1 ET0127 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF878219
VERSION BF878219.1 GI:12268349
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 369)
Dias Neto,E., Garcia Correa,R., Verjovsky-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

TITLE

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663

MEDLINE

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM2&t2=CM2-ET0127-181100-523-fl1&t3=2000-11-18&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 4

High quality sequence stop: 369.

FEATURES

source

Location/Qualifiers
1..369
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="ET0127"
/dev_stage="Adult"
/note="Organ: lung, tumor; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 82 a 116 c 73 g 98 t

ORIGIN

Query Match 3.3%; Score 34; DB 11; Length 369;

Best Local Similarity 100.0%; Pred. No. 2.4e-06;

Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 663 gtttcgccatgttgaccaggctgggtctcgaactc 696
|||||

Db 152 GTTTCGCCATGTTGACCAGGCTGGTCTCGAAGTC 185

RESULT 20

AQ346239

LOCUS

AQ346239 377 bp DNA

GSS

07-MAY-1999

DEFINITION RPCI11-116D16.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-116D16, DNA sequence.
ACCESSION AQ346239
VERSION AQ346239.1 GI:4171135
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 377)
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter,J.C.
TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready Map Building
JOURNAL Map Building
COMMENT Unpublished (1997)
Other_GSSs: RPCI11-116D16.TV
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.

FEATURES
source
1..377
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="GDB:7544247"
/db_xref="taxon:9606"
/clone="RPCI-11-116D16"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI; RPCI11 Human Male BAC Library"
77 a 114 c 79 g 107 t

BASE COUNT
ORIGIN

Query Match 3.3%; Score 34; DB 13; Length 377;
Best Local Similarity 100.0%; Pred. No. 2.4e-06;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 663 gtttcgccatgttgaccaggctggtctcgaaactc 696
|||||
Db 84 GTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 117

RESULT 21
AW864431
LOCUS AW864431 407 bp mRNA EST 22-MAY-2000
DEFINITION PM4-SN0016-030400-002-b12 SN0016 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW864431
VERSION AW864431.1 GI:7998390
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 407)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=PM4-SN0016-030400-002-b12&t3=2000-04-03&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 382.
Location/Qualifiers
1..407
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="SN0016"
/dev_stage="Adult"
/note="Organ: stomach_normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
73 a 95 c 122 g 117 t

FEATURES
source

Query Match 3.3%; Score 34; DB 10; Length 407;
Best Local Similarity 100.0%; Pred. No. 2.4e-06;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 663 gtttcgccatgttgaccaggctggtctcgaaactc 696
|||||
Db 200 GTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 233

RESULT 22
AQ093501/c
LOCUS AQ093501 412 bp DNA GSS 27-AUG-1998
DEFINITION HS_3018_B1_G09_T7 ClT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3018 Col=17 Row=N, DNA sequence.
ACCESSION AQ093501
VERSION AQ093501.1 GI:3464948
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 412)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 3018 row: N column: 17

Class: BAC ends
High quality sequence stop: 412.
Location/Qualifiers
1. .412
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="plate=3018 Col=17 Row=N"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBelolPAC11; BAC Clones in E-Coli DH10B"

BASE COUNT 140 a 94 c 95 g 83 t
ORIGIN

Query Match 3.3%; Score 34; DB 13; Length 412;
Best Local Similarity 100.0%; Pred. No. 2.4e-06;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 663 gtttcgcatgttgaccaggctggtctcgaactc 696
|||||
Db 210 GTTTCGCCATGTTGACCAGGCTGGTCTCGAATC 177

RESULT 23
AW958514
LOCUS AW958514 589 bp mRNA EST 01-JUN-2000
DEFINITION EST370584 MAGE resequences, MAGE Homo sapiens cDNA, mRNA sequence.
ACCESSION AW958514
VERSION AW958514.1 GI:8148198
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 589)
AUTHORS Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt ,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and Quackenbush,J.
TITLE Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray
JOURNAL Unpublished (2000)
COMMENT Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: johnq@tigr.org
Plate: 121
Seq primer: Reverse.
Location/Qualifiers
1. .589
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="MAGE resequences, MAGE"
/note="Vector: pBluescriptSKm"

BASE COUNT 153 a 131 c 131 g 174 t
ORIGIN

Query Match 3.3%; Score 34; DB 10; Length 589;
Best Local Similarity 100.0%; Pred. No. 2.4e-06;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 661 gggtttcgcatgttgaccaggctggtctcgaac 694
|||||
Db 355 GGGTTTCGCCATGTTGACCAGGCTGGTCTCGAAC 388

RESULT 24
BE296260/c
LOCUS BE296260 734 bp mRNA EST 20-JUL-2000
DEFINITION 601174838F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:3530265 5',

mRNA sequence.
BE296260
VERSION BE296260.1 GI:9179819
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 734)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LCM202 row: 1 column: 10
High quality sequence stop: 641.
Location/Qualifiers
1. .734
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3530265"
/clone_lib="NIH_MGC_17"
/tissue_type="rhabdomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: muscle; Vector: pOTB7; Site_1: EcoRI; Site_2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 183 a 175 c 204 g 172 t
ORIGIN

Query Match 3.3%; Score 34; DB 10; Length 734;
Best Local Similarity 100.0%; Pred. No. 2.4e-06;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 663 gtttcgcatgttgaccaggctggtctcgaactc 696
|||||
Db 561 GTTTCGCCATGTTGACCAGGCTGGTCTCGAATC 528

RESULT 25
AI051456/c
LOCUS AI051456 279 bp mRNA EST 28-AUG-1998
DEFINITION ow29b12.s1 Soares_parathyroid_tumor_NbHPA Homo sapiens cDNA clone IMAGE:1648223 3' similar to contains Alu repetitive element;; mRNA sequence.
ACCESSION AI051456
VERSION AI051456.1 GI:3306990
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 279)
AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo , Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 350 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham.

FEATURES
source

Location/Qualifiers
1. .279
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1648223"
/clone_lib="Soares_parathyroid_tumor_NbHPA"
/tissue_type="parathyroid tumor"
/dev_stage="adult"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: parathyroid gland; Vector: pT7T3D (Pharmacia
RI; 1st strand cDNA was primed with a Not I - oligo(dT)
primer
[5'-TGTTACCAATCTGAAGTGGGAGCGGCCGACCAATTTTTTTTTTTTTTTT
TTTTT-3'], double-stranded cDNA was size selected, ligated
to Eco RI adapters (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of a modified pT7T3
vector (Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid
adenomas was kindly provided by Dr. Stephen Marx, National
Institute of Diabetes and Digestive and Kidney Diseases,
NIH."

BASE COUNT 49 a 68 c 78 g 84 t
ORIGIN

Query Match 3.2%; Score 33; DB 10; Length 279;
Best Local Similarity 100.0%; Pred. No. 8.le-06;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 664 ttctgccatgttgaccaggctggtctcgaactc 696
|||||
Db 279 TTTCGCCATGTGACCAGGCTGGTCTCGAATC 247

RESULT 26
AW339622

LOCUS AW339622 298 bp mRNA EST 31-JAN-2000
DEFINITION hel5c05.x1 NCI_CGAP_CML1 Homo sapiens cDNA clone IMAGE:2919080 3'
similar to contains Alu repetitive element;contains element LTR8
repetitive element ;, mRNA sequence.

ACCESSION AW339622.1 GI:6836248
VERSION EST.
KEYWORDS human.
SOURCE Tumor Gene Index

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 298)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Elisabeth Paietta, Jonathan D. Licht, M.D.,
Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life
Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The
I.M.A.G.E. Consortium DNA Sequencing by: Washington University
Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 287.

FEATURES
source

Location/Qualifiers
1. .298
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2919080"
/clone_lib="NCI_CGAP_CML1"
/tissue_type="myeloid cells, 18 pooled CML cases, BCR/ABL
rearrangement positive, includes both chronic phase and
myeloid blast crisis"
/lab_host="DH10B"
/note="Organ: whole blood; Vector: pCMV-SPORT6; Site_1:
Sali; Site_2: NotI; Cloned unidirectionally. Primer:
Oligo dT. Library constructed by Life Technologies."

BASE COUNT 45 a 94 c 82 g 77 t
ORIGIN

Query Match 3.2%; Score 33; DB 10; Length 298;
Best Local Similarity 100.0%; Pred. No. 8e-06;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtaagacgggttcgccatgttgaccagg 682
|||||
Db 178 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGG 210

RESULT 27
AI244758

LOCUS AI244758 348 bp mRNA EST 28-JAN-1999
DEFINITION qj97gl2.x1 NCI_CGAP_Kid3 Homo sapiens cDNA clone IMAGE:1867462 3'
similar to contains Alu repetitive element;; mRNA sequence.

ACCESSION AI244758

VERSION AI244758.1 GI:3840155

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 348)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 304 Std Error: 0.00
Seq primer: -40UP from Gibco.

FEATURES
source

Location/Qualifiers
1. .348
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1867462"
/clone_lib="NCI_CGAP_Kid3"
/lab_host="DH10B"
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer,
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT7T3 vector. mRNA
source: 2 pooled kidneys. Library went through one round
of normalization. Library constructed by Bento Soares and
M. Fatima Bonaldo."

BASE COUNT 70 a 105 c 88 g 85 t
ORIGIN

Query Match 3.2%; Score 33; DB 10; Length 348;
Best Local Similarity 100.0%; Pred. No. 8e-06;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 664 ttctgccatgttgaccaggctgtctcgaactc 696
|||||
Db 170 TTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 202

RESULT 28
R98160 400 bp mRNA EST 11-SEP-1995
LOCUS
DEFINITION Yq66h03.r1 Soares fetal liver spleen lNFLS Homo sapiens cDNA clone
IMAGE:200789 5' similar to contains Alu repetitive element;contains
PTR5 repetitive element ;, mRNA sequence.
R98160
R98160.1 GI:983820
EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 400)
REFERENCE Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
AUTHORS M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
R., Williamson,A., Wohldmann,P. and Wilson,R.
The WashU-Merck EST Project
Contact: Wilson RK
Unpublished (1995)
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 644
High quality sequence stops: 267
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 644 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 267.
FEATURES
Location/Qualifiers
1..400
/organism="Homo sapiens"
/db_xref="GDB:3769839"
/db_xref="taxon:9606"
/clone="IMAGE:200789"
/clone_lib="Soares fetal liver spleen lNFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(df) primer
[5' AACTGGAGAAATTAAATAAGATCTTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 73 a 112 c 94 g 120 t 1 others
ORIGIN

Query Match 3.2%; Score 33; DB 11; Length 400;
Best Local Similarity 100.0%; Pred. No. 8e-06;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 669 ccatgttgaccaggctgtctcgaactcttgac 701

Db 195 CCATGTTGACCAGGCTGGTCTCGAACTCTTGAC 227
|||||
RESULT 29
AQ142429 405 bp DNA GSS 24-SEP-1998
LOCUS
DEFINITION HS_2244_A2_H04_MF CIT Approved Human Genomic Sperm Library D Homo
sapiens genomic clone Plate=2244 Col=8 Row=O, DNA sequence.
AQ142429
ACCESSION AQ142429.1 GI:3533082
VERSION
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 405)
REFERENCE Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
AUTHORS Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
99380589
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 2244 row: O column: 8
Class: BAC ends
High quality sequence stop: 405.
FEATURES
Location/Qualifiers
1..405
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=2244 Col=8 Row=O"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
E-Coli DH10B"
BASE COUNT 88 a 98 c 77 g 141 t 1 others
ORIGIN

Query Match 3.2%; Score 33; DB 13; Length 405;
Best Local Similarity 100.0%; Pred. No. 8e-06;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 664 ttctgccatgttgaccaggctgtctcgaactc 696
|||||
Db 118 TTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 150

RESULT 30
AQ455559 409 bp DNA GSS 21-APR-1999
LOCUS
DEFINITION HS_5157_B2_G10_T7A RPCI-11 Human Male BAC Library Homo sapiens
genomic clone Plate=733 Col=20 Row=N, DNA sequence.
AQ455559
ACCESSION AQ455559.1 GI:4596641
VERSION
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 409)
REFERENCE Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
AUTHORS Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Resear h Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu
Plate: 733 row: N column: 20
Seq primer: T7
Class: BAC ends
High quality sequence stop: 409.

FEATURES
source
Location/Qualifiers
1. .409
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=733 Col=20 Row=N"
/clone_lib="RPCI-11 Human Male BAC Library"
/sex="male"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at EcoRI sites"
BASE COUNT 118 a 105 c 96 g 86 t 4 others
ORIGIN
Query Match 3.2%; Score 33; DB 13; Length 409;
Best Local Similarity 100.0%; Pred. No. 8e-06;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 664 ttctgccatgttgaccaggctggtctcgaaactc 696
|||||
Db 277 TTTCGCCATGTTGACCAGCGTGGTCTCGAACTC 245
|||||

RESULT 31
M78087
LOCUS M78087 411 bp mRNA EST 26-MAY-1992
DEFINITION EST01677 Subtracted Hippocampus, Stratagene (cat. #936205) Homo sapiens cDNA clone HHCPE75 similar to Alu repetitive element, mRNA sequence.
ACCESSION M78087
VERSION M78087.1 GI:273824
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 411)
AUTHORS Adams,M.D., Dubnick,M., Kerlavage,A.R., Moreno,R., Kelley,J.M., Utterback,T.R., Nagle,J.W., Fields,C. and Venter,J.C.
TITLE Sequence identification of 2,375 human brain genes
JOURNAL Nature 355, 632-634 (1992)
MEDLINE 92168112
COMMENT Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
Seq primer: M13 Forward.

FEATURES
source
Location/Qualifiers
1. .411
/organism="Homo sapiens"
/db_xref="ATCC (inhost):78447"
/db_xref="GDB:D0S1086E"
/db_xref="taxon:9606"
/clone="HHCPE75"
/clone_lib="Subtracted Hippocampus, Stratagene (cat. #936205)"
/note="Vector: lambdaZAP-II; The hippocampus library (#4 above) was subtracted with a fibroblast cell line cDNA library (Stratagene cat. #936209; WI38 lung fibroblast cell line; oligo-dT + random primed cDNA synthesis; lambdaZAP-II vector, 1.0kb average insert size.) by the method of Sive & St. John (Nucl. Acids Res. 16:10937, 1988)."
BASE COUNT 99 a 114 c 93 g 105 t
ORIGIN

Query Match 3.2%; Score 33; DB 11; Length 411;
Best Local Similarity 100.0%; Pred. No. 8e-06;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 669 ccatgttgaccaggctggtctcgaaactcttgac 701
|||||
Db 63 CCATGTTGACCAGCGTGGTCTCGAACTCTTGAC 95
|||||
RESULT 32
AA457230/c
LOCUS AA457230 417 bp mRNA EST 06-JUN-1997
DEFINITION aa91c02.sl Stratagene fetal retina 937202 Homo sapiens cDNA clone IMAGE:838658 3' similar to contains Alu repetitive element;; mRNA sequence.
ACCESSION AA457230
VERSION AA457230.1 GI:2179950
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 417)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-Merck EST Project 1997
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Possible reversed clone: polyT not found
Seq primer: -41ml3 fwd. ET from Amersham
High quality sequence stop: 394.

FEATURES
source
Location/Qualifiers
1. .417
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:838658"
/clone_lib="Stratagene fetal retina 937202"
/sex="mixed"
/lab_host="SOLR (kanamycin resistant)"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Pooled retinal tissue. Average insert size: 1.0 Kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor sequence: 5' CTCGAGTTT TTT TTT TTT 3' "

BASE COUNT 129 a 95 c 83 g 110 t

ORIGIN

Query Match 3.2%; Score 33; DB 10; Length 417;
Best Local Similarity 100.0%; Pred. No. 7.9e-06;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 669 ccatgttgaccaggctggctcgaactcttgac 701
|||||

Db 347 CCATGTTGACCAGGCTGGTCTCGAACTCTTGAC 315

RESULT 33
BE146457/c 447 bp mRNA EST 21-JUN-2000
LOCUS BE146457 QV0-HT0215-041199-046-b01 HT0215 Homo sapiens cDNA, mRNA sequence.
DEFINITION BE146457
ACCESSION BE146457
VERSION BE146457.1 GI:8609181
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 447)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2=QV0-HT0215-041
199-046-b01&t3=1999-11-04&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 10
High quality sequence stop: 447.
High quality sequence stop: 447.
Location/Qualifiers
1..447
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0215"
/dev_stage="Adult"
/note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
low stringency conditions."
140 a 93 c 132 g 82 t

BASE COUNT 140 a 93 c 132 g 82 t

ORIGIN

Query Match 3.2%; Score 33; DB 10; Length 447;
Best Local Similarity 100.0%; Pred. No. 7.9e-06;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 669 ccatgttgaccaggctggctcgaactcttgac 701
|||||

Db 110 CCATGTTGACCAGGCTGGTCTCGAACTCTTGAC 78

RESULT 34
BF802026/c 450 bp mRNA EST 12-JAN-2001
LOCUS BF802026 CM1-CI0092-301000-503-h06 CI0092 Homo sapiens cDNA, mRNA sequence.
DEFINITION BF802026
ACCESSION BF802026.1 GI:12131015
VERSION BF802026.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 450)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM1&t2=CM1-CI0092-
301000-503-h06&t3=2000-10-30&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 49
High quality sequence stop: 450.
High quality sequence stop: 450.
Location/Qualifiers
1..450
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CI0092"
/dev_stage="Adult"
/note="Organ: colon_ins; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
low stringency conditions."
115 a 109 c 131 g 95 t

BASE COUNT 115 a 109 c 131 g 95 t

ORIGIN

Query Match 3.2%; Score 33; DB 11; Length 450;
Best Local Similarity 100.0%; Pred. No. 7.9e-06;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 669 ccatgttgaccaggctggctcgaactcttgac 701
|||||

Db 330 CCATGTTGACCAGGCTGGTCTCGAACTCTTGAC 298

RESULT 35
AA984191/c 453 bp mRNA EST 27-MAY-1998
LOCUS AA984191 am82b06.s1 Stratagene schizo brain S11 Homo sapiens cDNA clone
DEFINITION IMAGE:1629587 3' similar to contains Alu repetitive element;; mRNA
sequence.
ACCESSION AA984191

VERSION AA984191.1 GI:3162716
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 453)
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
White,Y., Wylie,T., Waterston,R. and Wilson,R.
WashU-NCI human EST Project
Unpublished (1997)
TITLE
JOURNAL
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Possible reversed clone: polyT not found
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 452.
Location/Qualifiers
1. .453
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1629587"
/clone_lib="Stratagene schizo brain s11"
/sex="male"
/tissue_type="schizophrenic brain S-11 frontal lobe"
/dev_stage="34 years old"
/lab_host="SOLR (kanamycin resistant)"
/note="Vector: Bluescript SK-; Site_1: EcoRI; Library
constructed from S-11 frontal lobe, male, 34 years old,
50% caucasian, 50% Aleutian. Schizophrenic suicide.
Random primed into EcoRI site of ZAP II Vector. Mass
excised. Custom library. Avg insert length 1.4kb.
Material obtained by Johnston N., Torrey, E.F., Volken R.,
and the Stanley Neuropathology Consortium - Analysis of
RNAs from the Brains of Individuals with Psychiatric
Diseases (Unpublished) Stanley Neurovirology Laboratory,
Johns Hopkins School of Medicine, Baltimore MD."

BASE COUNT 126 a 100 c 132 g 95 t
ORIGIN

Query Match 3.2%; Score 33; DB 10; Length 453;
Best Local Similarity 100.0%; Pred. No. 7.9e-06;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 664 ttctgcacatgttgaccaggctggtctcgaaactc 696
|||||
Db 220 TTTCGCCATGTTGACCCAGGCTGGTCTCGAACTC 188

RESULT 36
BF982842/c 679 bp mRNA EST 23-JAN-2001
LOCUS
DEFINITION 602305602F1 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:4396752 5',
mRNA sequence.
ACCESSION BF982842
VERSION BF982842.1 GI:12385654
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 679)
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10096 row: d column: 01
High quality sequence stop: 667.
Location/Qualifiers
1. .679
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4396752"
/clone_lib="NIH_MGC_88"
/tissue_type="duodenal adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: small intestine; Vector: pCMV-SPORT6;
Site_1: NotI; Site_2: SalI; Cloned unidirectionally;
oligo-dT primed. Average insert size 1.767 kb. Library
enriched for full-length clones and constructed by Life
Technologies. Note: this is a NIH_MGC Library."

BASE COUNT 208 a 135 c 160 g 176 t
ORIGIN

Query Match 3.2%; Score 33; DB 11; Length 679;
Best Local Similarity 100.0%; Pred. No. 7.8e-06;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 664 ttctgcacatgttgaccaggctggtctcgaaactc 696
|||||
Db 489 TTTCGCCATGTTGACCCAGGCTGGTCTCGAACTC 457

RESULT 37
AQ595304/c 712 bp DNA GSS 08-JUN-1999
LOCUS
DEFINITION HS_5448_A2_G08_T7A RPCI-11 Human Male BAC Library Homo sapiens
genomic clone Plate=1024 Col=16 Row=M, DNA sequence.
ACCESSION AQ595304
VERSION AQ595304.1 GI:5026890
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 712)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Resear h Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 1024 row: M column: 16
Seq primer: T7
Class: BAC ends

FEATURES	High quality sequence stop: 712.									
	Location/Qualifiers									
	1. .712									
	/organism="Homo sapiens"									
	/db_xref="taxon:9606"									
	/clone="Plate=1024 Col=16 Row=M"									
	/clone_lib="RPCI-11 Human Male BAC Library"									
	/sex="male"									
	/notes="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at EcoRI sites"									
	pBACe3.6 vector at EcoRI sites"									
BASE COUNT		223 a	141 c	157 g	166 t	25 others				
ORIGIN										
Query Match		3.2%; Score 33; DB 13; Length 712;								
Best Local Similarity		100.0%; Pred. No. 7.8e-06;								
Matches		33; Conservative	0; Mismatches	0; Indels	0; Gaps	0;				
QY		669	ccatgttgaccaggctggtctcgaaactcttgac	701						
Db		354	CCATGTTGACCAGGCTGGTCTCGAACTCTTGAC	322						
RESULT		38								
BG575508/c										
LOCUS										
DEFINITION		BG575508	859 bp	mRNA	EST	10-APR-2001				
		602598919F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:4707460 5', mRNA sequence.								
ACCESSION		BG575508								
VERSION		BG575508.1 GI:13583161								
KEYWORDS		EST.								
SOURCE		human.								
ORGANISM		Homo sapiens								
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.								
REFERENCE		1 (bases 1 to 859)								
AUTHORS		NIH-MGC http://mgc.nci.nih.gov/.								
TITLE		National Institutes of Health, Mammalian Gene Collection (MGC)								
JOURNAL		Unpublished (1999)								
COMMENT		Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: DCTD/DTP CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM10577 row: f column: 05 High quality sequence stop: 812.								
FEATURES		Location/Qualifiers								
source		1. .859								
		/organism="Homo sapiens"								
		/db_xref="taxon:9606"								
		/clone="IMAGE:4707460"								
		/clone_lib="NIH_MGC_87"								
		/tissue_type="mammary adenocarcinoma, cell line"								
		/lab_host="DH10B (phage-resistant)"								
		/note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.383 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."								
BASE COUNT		260 a	158 c	197 g	243 t	1 others				
ORIGIN										
Query Match		3.2%; Score 33; DB 11; Length 859;								
Best Local Similarity		100.0%; Pred. No. 7.8e-06;								
Matches		33; Conservative	0; Mismatches	0; Indels	0; Gaps	0;				

AUTHORS NCI/NIDR-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute / National Institute of Dental Research,
Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: John F. Ensley, M.D., Chidehanok Leethanakul,
D.D.S.

cDNA Library Preparation: Krizman and Leethanakul Laboratories
cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov

Seq primer: -40UP from Gibco
High quality sequence stop: 211.

FEATURES

source
1. .235
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4318471"
/clone_lib="NCI-CGAP_HN14"
/tissue_type="hyperplasia of squamous epithelium"
/lab_host="DH10B"

/note="Organ: tongue; Vector: pAMP10; mRNA made from
tongue epithelium, cDNA made by oligo-dT priming.
Non-directionally cloned into UDG sites. Size-selected on
agarose gel, average insert size 500 bp. Primary library.
cDNA Library Preparation: David B. Krizman, Ph.D.
REFERENCE: Krizman et al. (1996) Cancer Research
56:5380-5383."

BASE COUNT 44 a 57 c 63 g 71 t
ORIGIN

Query Match

Best Local Similarity 3.1%; Score 32; DB 11; Length 235;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 681 ggctggtctcgaaactcttgacgtcaagtgatc 712
|||||
Db 198 GGCTGGTCTCGAACTCTTGACGTCAGTGATC 229

RESULT 41

AQ095034 392 bp DNA GSS 27-AUG-1998
LOCUS HS_3028_A2_G11_MF CIT Approved Human Genomic Sperm Library D Homo
DEFINITION sapiens genomic clone Plate=3028 Col=22 Row=M, DNA sequence.
ACCESSION AQ095034
VERSION AQ095034.1 GI:3466451
KEYWORDS GSS.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 392)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.

Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
99380589

COMMENT

Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 3028 row: M column: 22

Class: BAC ends
High quality sequence stop: 392.
Location/Qualifiers
1. .392
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=3028 Col=22 Row=M"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
E-Coli DH10B"

BASE COUNT 85 a 80 c 85 g 141 t 1 others
ORIGIN

Query Match 3.1%; Score 32; DB 13; Length 392;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 668 gccatgttgaccaggctggtctcgaaactcttg 699
|||||
Db 361 GCCATGTTGACCAGGCTGGTCTCGAACTCTTG 392

RESULT 42

AG002404/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. .616

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="21"

/clone="149C3X10"

/map="21q"

BASE COUNT 171 a 125 c 161 g 155 t 4 others

ORIGIN

Query Match 3.1%; Score 32; DB 13; Length 616;

Best Local Similarity 100.0%; Pred. No. 2.6e-05;

Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 658 acgggggttcgccatgttgaccaggctggtct 689

|||||

Db 140 ACGGGGTTTCGCCATGTGACCAGCGTGGTCT 109

RESULT 43

AG002403/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

Contact: Mahairas GG, Wallace JC, Hood L

High Throughput Sequencing Center

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Sequence Tagged Connector

Plate: 3028 row: M column: 22

```
ACCESSION AG002403
VERSION AG002403.1 GI:2627055
KEYWORDS GSS.
SOURCE Homo sapiens DNA, clone:149C3X10.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 658)
AUTHORS Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-1997) to the DDBJ/EMBL/GenBank databases.
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory; Kitasato 1-15-1, Sagamihara 228, Japan
(E-mail:hattori@hgc.ims.u-tokyo.ac.jp, Tel:0427-78-9732,
Fax:0427-78-9561)
REFERENCE 2 (bases 1 to 658)
AUTHORS Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
TITLE Homo sapiens genomic DNA, chromosome 21q
JOURNAL Published Only in DataBase (1997) In press
FEATURES             Location/Qualifiers
     source           1..658
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /clone="149C3X10"
                     /map="21q"
BASE COUNT 188 a 126 c 168 g 168 t 8 others
ORIGIN

Query Match      3.1%; Score 32; DB 13; Length 658;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 658 acgggggttcgccatgttgaccaggctggtct 689
|||||
Db 169 ACGGGGTTTCGCCATGTTGACCAAGGCTGGTCT 138

RESULT 44
AG002384
LOCUS AG002384 678 bp DNA GSS 06-FEB-1999
DEFINITION Homo sapiens genomic DNA, 21q region, clone: 149C3X10, genomic
survey sequence.
ACCESSION AG002384
VERSION AG002384.1 GI:2627036
KEYWORDS GSS.
SOURCE Homo sapiens DNA, clone:149C3X10.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 678)
AUTHORS Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-1997) to the DDBJ/EMBL/GenBank databases.
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory; Kitasato 1-15-1, Sagamihara 228, Japan
(E-mail:hattori@hgc.ims.u-tokyo.ac.jp, Tel:0427-78-9732,
Fax:0427-78-9561)
REFERENCE 2 (bases 1 to 678)
AUTHORS Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
TITLE Homo sapiens genomic DNA, chromosome 21q
JOURNAL Published Only in DataBase (1997) In press
FEATURES             Location/Qualifiers
     source           1..678
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /clone="149C3X10"
                     /map="21q"
BASE COUNT 191 a 160 c 128 g 195 t 4 others
ORIGIN
```

```
Query Match      3.1%; Score 32; DB 13; Length 678;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 658 acgggggttcgccatgttgaccaggctggtct 689
|||||
Db 126 ACGGGGTTTCGCCATGTTGACCAAGGCTGGTCT 157

RESULT 45
AL048065
LOCUS AL048065 687 bp mRNA EST 01-MAR-2000
DEFINITION DKFZp586G1523_r1 586 (synonym: hutel) Homo sapiens cDNA clone
DKFZp586G1523, mRNA sequence.
ACCESSION AL048065
VERSION AL048065.1 GI:4728898
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 687)
AUTHORS Ansorge,W., Wirkner,U., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
TITLE EST (Ansorge, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Ansorge W
MIPS
Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
s1 sequence also available.
This clone (DKFZp586G1523) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES             Location/Qualifiers
     source           1..687
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DKFZp586G1523"
                     /clone_lib="586 (synonym: hutel)"
                     /tissue_type="uterus"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Vector: pSport1; Site_1: NotI; Site_2: SalI/MluI"
BASE COUNT 192 a 139 c 137 g 219 t
ORIGIN

Query Match      3.1%; Score 32; DB 10; Length 687;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 ccattgtgaccaggctggtctcgaaactcttga 700
|||||
Db 279 CCATGTTGACCAAGGCTGGTCTCGAACTCTTGA 310

RESULT 46
AG002385
LOCUS AG002385 691 bp DNA GSS 06-FEB-1999
DEFINITION Homo sapiens genomic DNA, 21q region, clone: 149C3X10, genomic
survey sequence.
ACCESSION AG002385
VERSION AG002385.1 GI:2627037
KEYWORDS GSS.
SOURCE Homo sapiens DNA, clone:149C3X10.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 691)
TITLE Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
JOURNAL Direct Submission
Submitted (18-NOV-1997) to the DDBJ/EMBL/GenBank databases.
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory; Kitasato 1-15-1, Sagamihara 228, Japan
(E-mail:hattori@hc.ims.u-tokyo.ac.jp, Tel:0427-78-9732,
Fax:0427-78-9561)
REFERENCE 2 (bases 1 to 691)
AUTHORS Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
TITLE Homo sapiens genomic DNA, chromosome 21q
JOURNAL Published Only in DataBase (1997) In press
FEATURES Location/Qualifiers
source
1. .691
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/clone="149C3X10"
/map="21q"

BASE COUNT 204 a 163 c 124 g 200 t
ORIGIN

Query Match 3.1%; Score 32; DB 13; Length 691;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 658 acgggggttcgccatgttgaccagctgtgtct 689
|||||
Db 343 ACGGGGTTTCGCCATGTGACCGCTGGTCT 374

RESULT 47
AQ742378
LOCUS AQ742378 853 bp DNA GSS 16-JUL-1999
DEFINITION HS_5382_B1_C09_T7A RPCI-11 Human Male BAC Library Homo sapiens
genomic clone Plate=958 Col=17 Row=F, DNA sequence.
ACCESSION AQ742378
VERSION AQ742378.1 GI:5519900
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 853)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Resear h Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 958 row: F column: 17
Seq primer: T7
Class: BAC ends
High quality sequence stop: 853.

FEATURES
source
1. .853
/organism="Homo sapiens"

/db_xref="taxon:9606"
/clone="Plate=958 Col=17 Row=F"
/clone_lib="RPCI-11 Human Male BAC Library"
/sex="male"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBACe3.6 vector at EcoRI sites"
BASE COUNT 182 a 220 c 181 g 270 t
ORIGIN

Query Match 3.1%; Score 32; DB 13; Length 853;
Best Local Similarity 100.0%; Pred. No. 2.5e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 670 catgttgaccagctggtctcgaactcttgac 701
|||||
Db 477 CATGTTGACCGAGCTGGTCTCGAACTCTTGAC 508

RESULT 48
AA640211
LOCUS AA640211 248 bp mRNA EST 23-OCT-1997
DEFINITION nn95h08.s1 NCI_CGAP_Phe1 Homo sapiens cDNA clone IMAGE:1098975 3'
similar to contains Alu repetitive element;; mRNA sequence.
ACCESSION AA640211
VERSION AA640211.1 GI:2563990
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 248)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: W. Marston Linehan, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Stratagene, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
www-bio.lnl.gov/bbrp/image/image.html
Seq primer: -40ml3 fwd. ET from Amersham.

FEATURES
source
1. .248
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1098975"
/clone_lib="NCI_CGAP_Phe1"
/tissue_type="pheochromocytoma"
/lab_host="SOLR (kanamycin resistant)"
/note="Vector: Bluescript SK-; Site_1: EcoRI; Site_2: XhoI
; Cloned unidirectionally. Primer: Oligo dt.
Pheochromocytoma. 5' adaptor sequence: 5' GAATTCGGCAGAG
3' 3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'
Average insert size: 1.3 kb."
BASE COUNT 46 a 69 c 63 g 70 t
ORIGIN

Query Match 3.0%; Score 31; DB 10; Length 248;
Best Local Similarity 100.0%; Pred. No. 8.6e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 660 ggggttcgccatgttgaccagctggtctc 690
|||||

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2002, 11:02:15 ; Search time 40.16 Seconds
(without alignments)
581.003 Million cell updates/sec

Title: US-09-671-050-12
Perfect score: 1662
Sequence: 1 MEKYEKLAKTGEYSYGVVFK.....RKARNEGRRRRQQVLP LKS 315

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_ll01.*
1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT.*
2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1662	100.0	315	22 AAE00495	Human kinase #6.
2	1638	98.6	324	22 AAE00492	Human kinase #3.
3	1636	98.4	347	22 AAE00494	Human kinase #5.
4	1619	97.4	360	22 AAU03525	Human protein kina
5	1612	97.0	356	22 AAE00491	Human kinase #2.
6	1296.5	78.0	296	22 AAB65643	Novel protein kina
7	1149	69.1	247	22 AAB65642	Novel protein kina
8	967.5	58.2	566	21 AAY90724	Rabbit KKIAMRE kin
9	791	47.6	187	22 AAE00490	Human kinase #1.
10	791	47.6	198	22 AAE00493	Human kinase #4.
11	784.5	47.2	455	21 AAY70126	Human lost in leuk

12	784.5	47.2	591	22 AAU03524	Human protein kina
13	763.5	45.9	457	21 AAY70125	Rat lost in leukae
14	763.5	45.9	505	21 AAY70124	Rat lost in leukae
15	656	39.5	534	22 AAB65645	Novel protein kina
16	581	35.0	333	22 AAM39276	Human polypeptide
17	574	34.5	352	22 AAM41062	Human polypeptide
18	570	34.3	298	20 AAW95689	Homo sapiens Cdc2
19	570	34.3	544	19 AAW85028	CDK2-green fluores
20	570	34.3	544	19 AAW85029	Green fluorescent
21	568.5	34.2	294	20 AAW95690	Oryza sativa Cdc2
22	566.5	34.1	294	21 AAG33356	Zea mays protein f
23	565.5	34.0	294	21 AAG40809	Zea mays protein f
24	561.5	33.8	294	21 AAG11214	Arabidopsis thalia
25	561.5	33.8	294	21 AAG54042	Arabidopsis thalia
26	561.5	33.8	350	21 AAG11213	Arabidopsis thalia
27	557	33.5	297	20 AAW95687	Rattus norvegicus
28	556	33.5	297	20 AAW95688	Homo sapiens Cdc2
29	555.5	33.4	300	20 AAW95679	Pneumocystis carin
30	555	33.4	270	20 AAY43932	Human protein kina
31	554	33.3	298	21 AAY52184	Cyclin-dependent k
32	544	32.7	292	15 AAR53400	Cyclin dependent p
33	542	32.6	292	21 AAB35797	Protein involved i
34	539.5	32.5	294	22 AAB97609	Tomato cyclin-depe
35	518.5	31.2	297	20 AAW95686	Schizosaccharomyce
36	511	30.7	274	20 AAY43931	Fission yeast prot
37	510.5	30.7	311	21 AAG29618	Arabidopsis thalia
38	507	30.5	348	21 AAG14009	Arabidopsis thalia
39	507	30.5	357	21 AAG14008	Arabidopsis thalia
40	505.5	30.4	317	18 AAW01559	CDK1 protein. Can
41	505.5	30.4	317	20 AAW95682	Candida albicans C
42	505.5	30.4	317	21 AAY15116	Candida cyclin dep
43	505	30.4	323	20 AAW95685	EmERICELLA nidulan
44	500	30.1	346	21 AAY90243	Human cyclin depen
45	498	30.0	306	20 AAY49160	Soybean cyclin-dep

ALIGNMENTS

RESULT 1
AAE00495
ID AAE00495 standard; Protein; 315 AA.
XX
AC AAE00495;
XX
DT 19-JUN-2001 (first entry)
XX
DE Human kinase #6.
XX
KW Human; kinase; gene therapy; bioreactor; mental disorder;
KW biological disorder.
XX
OS Homo sapiens.
XX
PN WO200123579-A1.
XX
PD 05-APR-2001.
XX
PF 27-SEP-2000; 2000WO-US266621.
XX
PR 28-SEP-1999; 99US-0156511.
XX
(LEXI-) LEXICON GENETICS INC.
XX
PI Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
XX
DR WPI; 2001-266166/27.
DR N-PSDB; AAD03817.
XX
PT New isolated human kinase polynucleotide useful for generating
PT antibodies, as reagents in diagnostic assays and for screening for
PT compounds useful for treating mental, biological or medical diseases
XX

Db 301 egrnrrrqv 310
|||||
RESULT 3
AAE00494
ID AAE00494 standard; Protein; 347 AA.
XX AC AAE00494;
XX 19-JUN-2001 (first entry)
DT Human kinase #5.
DE Human; kinase; gene therapy; bioreactor; mental disorder;
XX biological disorder.
KW Homo sapiens.
KW WO200123579-A1.
XX 05-APR-2001.
PD 27-SEP-2000; 2000WO-US26621.
PF 28-SEP-1999; 99US-0156511.
PR (LEXI-) LEXICON GENETICS INC.
XX Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
PI WPI; 2001-266166/27.
XX N-PSDB; AAD03816.
DR New isolated human kinase polynucleotide useful for generating
XX antibodies, as reagents in diagnostic assays and for screening for
PT compounds useful for treating mental, biological or medical diseases -
PT Claim 2; Page 32-33; 38pp; English.
XX The present sequence is novel human protein
CC (NHP) known as human kinase. The human kinases share structural
CC similarity with animal kinases, more particularly serine or
CC threonine protein kinases. Human kinase cDNA is useful for the
CC detection of mutant human kinase for the diagnosis of disease,
CC and also as a therapeutic. It is useful for screening drugs
CC effective in the treatment of symptomatic or phenotypic
CC manifestations perturbing the normal function of NHP in the
CC body. The NHP nucleotide sequences are useful for generation of
CC antibodies, as reagents in diagnostic assays, for the
CC identification of other cellular gene products related to human
CC kinases, and as reagents in assays for screening compounds that
CC are useful for treating mental, biological or medical disorders.
CC NHP oligonucleotides are used as probes. The labelled NHP probes
CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
XX modulation of NHP expression.
SQ Sequence 347 AA;
Query Match 98.4%; Score 1636; DB 22; Length 347;
Best Local Similarity 90.8%; Pred. No. 2.6e-166;
Matches 315; Conservative 0; Mismatches 0; Indels 32; Gaps 1;
Qy 1 MEKYELAKTGEYSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKIALREIRMLKQLKH 60
|||||

Db 1 mekyelaktgeysygvvfkcrnktsgqvavkkfveseddpvkvkialreirmlkqlkh 60
Qy 61 PNLVNLIEVFRRRKRKMHVFEYCDHTLLNELERNPNGVADGVIKSVLWQTLQALNFCIH 120
Db 61 pnlvnlievfrkrkmlhvfeycdhtllnelernpngvadgviksvlwtlqalnfcih 120
Qy 121 NCIHRIKPENILITKQGIKICDFGAQIILIPGDAYTDYVATRWYRAPPELLVGDYQYGS 180
Db 121 ncihrdikpenilitkglikicdfgaqiilip9daytdyvatrwyrappelvldtqygs 180
Qy 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRT----- 217
Db 181 svdiwaigcvfaelltgqplwpgksdvdqlyliirtlvetgrhvdqaglelltssppa 240
Qy 218 -----GKLIPRHQSIFKSNGFHGISIPEPEDMETLEKFSVDVHPVALNFMKGCLK 268
Db 241 vasqsagitgkliprhqsfksgnfghgisipepedmetleekfsdvhpvalnfmkgclk 300
Qy 269 MNPDDRLLTCSOLLLESSYFDSFQEAQIKRKARNEGRRRQVLPKLS 315
Db 301 mnpddrltcsqlllessyfdsfqaqikrkarnegrnrrrqvlpkls 347
RESULT 4
AAU03525
ID AAU03525 standard; Protein; 360 AA.
XX AC AAU03525;
XX 12-SEP-2001 (first entry)
DT Human protein kinase #25.
DE Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
XX metabolic disorder; immune related disease; neurological disorder;
KW neurodegenerative disorder; inflammatory disorder; infectious disease;
KW reproductive disorder.
XX Homo sapiens.
XX WO200138503-A2.
PN 31-MAY-2001.
XX 22-NOV-2000; 2000WO-US32085.
XX 24-NOV-1999; 99US-0167482.
PR (SUGE-) SUGEN INC.
PA Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
XX Flanagan P, Clary D;
PI WPI; 2001-343950/36.
DR N-PSDB; AAS06725.
XX Nucleic acids encoding human kinase polypeptides, useful for preventing
PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT neuronal-associated diseases, and microbial infections -
XX Claim 7; Figure 2; 433pp; English.
XX AAU03501-AAU03557 represent novel human protein kinases #1-57. The
CC novel protein kinases have been identified as members of the tyrosine
CC or serine/threonine kinase (PTK and STK) families. The polynucleotides
CC encoding protein kinases and the polypeptides may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate kinase expression. For example, they may be used to treat
CC cancers (especially cancers of haematopoietic origin), cardiovascular
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC immune related diseases (e.g. rheumatoid arthritis), neurological
CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC parkinson's disease), inflammatory disorders (e.g. asthma), infectious

KW inflammatory pelvic disease; multiple sclerosis; psoriasis.
XX Mus musculus.
OS WO200073469-A2.
XX 07-DEC-2000.
XX 26-MAY-2000; 2000WO-US14842.
PF 28-MAY-1999; 99US-0136503.
PR (SUGE-) SUGEN INC.
XX Plowman GD, Martinez R, Whyte D, Sudersanam S;
PI WPI; 2001-032161/04.
DR N-PSDB; AAF44670.
XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and
PT treating immune-related diseases and disorders, cardiovascular disease,
PT neurodegenerative diseases and/or cancers -
XX Claim 10; Fig 1; 310pp; English.
PS The present sequence is a novel protein kinase. The novel protein kinases
XX and the nucleic acids that encode them may be used in the treatment and
CC diagnosis of diseases associated with inappropriate kinase expression
CC such as immune-related diseases and disorders, cardiovascular disease,
CC neurodegenerative diseases and/or cancers. The nucleic acids and
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays. The kinase polypeptides may be used as antigens in the production
CC of antibodies of kinase expression and activity. Anti-kinase antibodies
CC and kinase antagonists may also be used to down regulate kinase
CC expression and activity. Diseases related to kinase expression and
CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
CC disorders, complications of organ transplantation, myocardial infarction,
CC immune disorders, cardiomyopathies, strokes, renal failure,
CC oxidative-stress related disorders, chronic inflammatory bowel disease,
CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
CC reproductive disorders.
XX Sequence 296 AA;
SQ Query Match 78.0%; Score 1296.5; DB 22; Length 296;
Best Local Similarity 90.0%; Pred. No. 4.2e-130;
Matches 243; Conservative 19; Mismatches 5; Indels 3; Gaps 2;
Qy 46 KIALREIRMLKQLKHPNLVNLIEVFRKRKMHLVFEYCDHTLLNELRNPNGVADGVIKS 105
Db 1 kialreirmlk-lkhpnlvnlievfrkrkmhlvfeycdhtllnelrnpngvsdgviks 59
Qy 106 VLWQTLQALNFCCHIHRDIKPENILITKQGIKICDFGFAQILIPGDAYTDYVATRW 165
Db 60 vlwqtlqalnfcchkncihrdvxpennilitkgmikicdfgfarilipgdaydyvatrw 119
Qy 166 YRAPELLVGDTOYGSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIRTLGLIPRHQ 225
Db 120 yrapellvgdtkygssvdvwavgcvfaelltgqplwpgksdvdqyliirtlglkhiprhq 179
Qy 226 SIFKSNGFPHGISIPEPEDMETLEEKFSVDHPVALNFMKGCLKMNPDDRLLTCSQLESSEY 285
Db 180 sifrsnqfgrgisipepedmetleekfsnvqpvalsfmkgclkmnpdrrltaqildsay 239
Qy 286 FDSFQEAQIKRKARNEGRRNRQ--QVLPL 313
Db 240 fesfqedqmrkarsegrsrrrqngllpl 269
RESULT 7
AAB65642

ID AAB65642 standard; Protein; 247 AA.
XX AAB65642;
AC 27-MAR-2001 (first entry)
XX Novel protein kinase, SEQ ID NO: 169.
DE Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
XX immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic;
KW dermatological; antidiabetic; antiinfertility; gene therapy; vaccine;
KW immune disorder; cardiovascular disease; neurodegenerative disease;
KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
KW inflammatory pelvic disease; multiple sclerosis; psoriasis.
XX Homo sapiens.
OS WO200073469-A2.
XX 07-DEC-2000.
XX 26-MAY-2000; 2000WO-US14842.
PF 28-MAY-1999; 99US-0136503.
XX (SUGE-) SUGEN INC.
PA Plowman GD, Martinez R, Whyte D, Sudersanam S;
PI WPI; 2001-032161/04.
DR N-PSDB; AAF44669.
XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and
PT treating immune-related diseases and disorders, cardiovascular disease,
PT neurodegenerative diseases and/or cancers -
XX Claim 10; Fig 1; 310pp; English.
PS The present sequence is a novel protein kinase. The novel protein kinases
XX and the nucleic acids that encode them may be used in the treatment and
CC diagnosis of diseases associated with inappropriate kinase expression
CC such as immune-related diseases and disorders, cardiovascular disease,
CC neurodegenerative diseases and/or cancers. The nucleic acids and
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays. The kinase polypeptides may be used as antigens in the production
CC of antibodies of kinase expression and activity. Anti-kinase antibodies
CC and kinase antagonists may also be used to down regulate kinase
CC expression and activity. Diseases related to kinase expression and
CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
CC disorders, complications of organ transplantation, myocardial infarction,
CC immune disorders, cardiomyopathies, strokes, renal failure,
CC oxidative-stress related disorders, chronic inflammatory bowel disease,
CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
CC reproductive disorders.
XX Sequence 247 AA;
SQ Query Match 69.1%; Score 1149; DB 22; Length 247;
Best Local Similarity 98.6%; Pred. NO. 1.9e-114;
Matches 217; Conservative 1; Mismatches 0; Indels 2; Gaps 1;
Qy 96 NGVADGVIKSVLWQTLQALNFCCHIHRDIKPENILITKQGIKICDFGFAQILIPGD 155
Db 1 ngvadgviksvlwqtlqalnfcchihrdikpenilitkggikicdfgfaqiliipgd 60
Qy 156 AYTDYVATRWYRAPELLVGDTOYGSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIR 215
Db 61 aytdivatrwyrapellvgdtqygssvdiwaigcvfaelltggplwpgksdvdqyliir 120
Qy 216 TLGKLIPRHOSIFKSNGFPHGISIPEPEDMETLEKFSVDHPVALNFMKGCLKMNPDDRLL 275

Db 121 tlgkllprhqsifksngffhghisipepedmetleekfsdvhpvalnfmkgclkmnpddr1 180

QY 276 TCSQLLESSYFDSFOEAQIKRKARNEGRNRRRQ--QVLPL 313
||||| 177; Conservative 67; Mismatches 65; Indels 15; Gaps 5;

Db 181 tcsqllessyfdsfgeaqikrkarnegnrnrrrqnqlp1 220

RESULT 8

AAAY90724

ID AAY90724 standard; Protein; 566 AA.

XX AC AAY90724;

XX DT 15-AUG-2000 (first entry)

XX DE Rabbit KKIAMRE kinase SEQ ID NO:4.

XX KW Rabbit; KKIAMRE kinase; learning-induced kinase; learning; memory;

XX KW cdc2-related kinase; brain; gene therapy; genetic disorder; detection; identification.

XX OS Oryctolagus cuniculus.

XX XX WO2000020567-A2.

PN 13-APR-2000.

PD 01-OCT-1999; 99WO-US23010.

PF 02-OCT-1998; 98US-0102906.

XX (UYSC-) UNIV SOUTHERN CALIFORNIA.

PA Thompson RF, Gomi H, Sun W;

XX WPI; 2000-328932/28.

DR N-PSDB; AAA29745.

XX Novel learning induced kinase polynucleotides and polypeptides, useful for the analysis of learning and memory, and for gene therapy - Claim 1; Fig 4; 64pp; English.

XX The present sequence represents a learning-induced kinase, designated KKIAMRE kinase, which is isolated from rabbit brain tissue. KKIAMRE kinase is a cdc2-related kinase. The KKIAMRE kinase polynucleotides can be used to express recombinant protein for analysis, characterisation or therapeutic use, as markers for tissues in which the protein is preferentially expressed, as molecular weight markers on Southern gels, as chromosome markers or tags, to compare endogenous DNA sequences in patients to identify potential genetic disorders, as probes to hybridise and discover novel related sequences, as a source of PCR primers, and as an antigen to induce anti-DNA antibodies. The polypeptides can be used in assay to discover biological activity, to raise antibodies, as tissue markers, and to isolate correlative receptors or ligands. The polynucleotides may also be used for gene therapy for the treatment of disorders which are mediated by KKIAMRE kinase.

XX .

SQ Sequence 566 AA;

Query Match

Best Local Similarity 58.2%; Score 967.5; DB 21; Length 566;

Matches 177; Conservative 67; Mismatches 65; Indels 15; Gaps 5;

QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
||||| 177; Conservative 67; Mismatches 65; Indels 15; Gaps 5;

Db 1 mekyenlglvgeysygmvmkcrnkdsgrivaikkflesdddkmvmkiamreikllkqlrh 60

QY 61 PNLVNLIEVFRKKMKHLVFEYCDHTLLNELERNPNGVADGVKSVLWQTLQALNFCCHI 120
||||| 177; Conservative 67; Mismatches 65; Indels 15; Gaps 5;

Db 61 enlvnllevckkkrrwylvfefvdhtlddlelfpnglddqvqkylfqiingigfchsh 120

QY 121 NCIHRIKIPENILITKQGIKICDFGFAQL-IPGDAYTDYVATRWYRAPELLVGDTOYG 179
||||| 177; Conservative 67; Mismatches 65; Indels 15; Gaps 5;

Db 121 nihrdikpenilvsqsgvkvkcdfgfartlaapgevydyvatrwyrapellvgdvkyg 180

QY 180 SSVDIWAIGCVFAELITGQPLWPGKSDVDQYLIIRTGLKLIIPRHOSIFKSNGFHGISI 239
:||||| 177; Conservative 67; Mismatches 65; Indels 15; Gaps 5;

Db 181 kavdvwaigclvtelnmgeplfpgdsdidqlylimrcnlgnlprhqelfyknppvfvagvrl 240

QY 240 PEPEDMETLEEKFSDVHPVALNFMKGCLKMNPDDRLTCSQLLESSYF--DSF-----QEA 292
||:| 177; Conservative 67; Mismatches 65; Indels 15; Gaps 5;

Db 241 peikesepkerrypklsevvldlakkclhvdpkrfpcaellhhdfqmdgfaerfsqel 300

QY 293 QIK--RKARN-----EGRNRRRQQ 309
|:| 177; Conservative 67; Mismatches 65; Indels 15; Gaps 5;

Db 301 qmkvqkdarnislskksqnrkke 324

RESULT 9

AAE00490

ID AAE00490 standard; Protein; 187 AA.

XX AC AAE00490;

XX DT 19-JUN-2001 (first entry)

XX DE Human kinase #1.

XX KW Human; kinase; gene therapy; bioreactor; mental disorder; biological disorder.

XX OS Homo sapiens.

XX PN WO200123579-A1.

XX PD 05-APR-2001.

XX PF 27-SEP-2000; 2000WO-US26621.

XX PR 28-SEP-1999; 99US-0156511.

XX (LEXI-) LEXICON GENETICS INC.

XX Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT; WPI; 2001-266166/27.

DR N-PSDB; AAD03812, AAD03818.

XX New isolated human kinase polynucleotide useful for generating antibodies, as reagents in diagnostic assays and for screening for compounds useful for treating mental, biological or medical diseases - Disclosure; Page 27-28; 38pp; English.

XX The present sequence is novel human protein (NHP) known as human kinase. The human kinases share structural similarity with animal kinases, more particularly serine or threonine protein kinases. Human kinase cDNA is useful for the detection of mutant human kinase for the diagnosis of disease, and also as a therapeutic. It is useful for screening drugs effective in the treatment of symptomatic or phenotypic manifestations perturbing the normal function of NHP in the body. The NHP nucleotide sequences are useful for generation of antibodies, as reagents in diagnostic assays, for the identification of other cellular gene products related to human kinases, and as reagents in assays for screening compounds that are useful for treating mental, biological or medical disorders. NHP oligonucleotides are used as probes. The labelled NHP probes are useful for screening human genomic library for identifying polymorphisms and as primers in amplification assays to detect mutations within the exons, introns and splice sites that can be used in diagnostics and pharmacogenomics. Nucleotide construct encoding NHP products are used to genetically engineer cells in vivo that functions as bioreactors in the body delivering a

CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.

XX Sequence 187 AA;

Query Match 47.6%; Score 791; DB 22; Length 187;
Best Local Similarity 100.0%; Pred. No. 2.5e-76;
Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKYEKLAKTGEGSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 61 PNLVNLIEVFRKRKMHVLVFEYCDHTLLNELERNPNGVADGVIKSVLWQTIQALNFCIH 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 121 NCIHRIKIPENILITKQGIITKICDFGFAQIL 151
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 10
AAE00493
ID AAE00493 standard; Protein; 198 AA.

XX AC AAE00493;
XX 19-JUN-2001 (first entry)
XX Human kinase #4.
XX Human; kinase; gene therapy; bioreactor; mental disorder;
KW biological disorder.

XX OS Homo sapiens.
XX WO200123579-A1.
XX 05-APR-2001.
XX 27-SEP-2000; 2000WO-US26621.
XX 28-SEP-1999; 99US-0156511.
XX (LEXI-) LEXICON GENETICS INC.
XX Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
PI WPI; 2001-266166/27.
XX N-PSDB; AAD03815.

XX New isolated human kinase polynucleotide useful for generating
PT antibodies, as reagents in diagnostic assays and for screening for
PT compounds useful for treating mental, biological or medical diseases
XX Disclosure; Page 31; 38pp; English.

XX The present sequence is novel human protein
CC (NHP) known as human kinase. The human kinases share structural
CC similarity with animal kinases, more particularly serine or
CC threonine protein kinases. Human kinase cDNA is useful for the
CC detection of mutant human kinase for the diagnosis of disease,
CC and also as a therapeutic. It is useful for screening drugs
CC effective in the treatment of symptomatic or phenotypic
CC manifestations perturbing the normal function of NHP in the
CC body. The NHP nucleotide sequences are useful for generation of
CC antibodies, as reagents in diagnostic assays, for the
CC identification of other cellular gene products related to human
CC kinases, and as reagents in assays for screening compounds that
CC are useful for treating mental, biological or medical disorders.

CC NHP oligonucleotides are used as probes. The labelled NHP probes
CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.
XX Sequence 198 AA;

Query Match 47.6%; Score 791; DB 22; Length 198;
Best Local Similarity 100.0%; Pred. No. 2.7e-76;
Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKYEKLAKTGEGSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 61 PNLVNLIEVFRKRKMHVLVFEYCDHTLLNELERNPNGVADGVIKSVLWQTIQALNFCIH 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 121 NCIHRIKIPENILITKQGIITKICDFGFAQIL 151
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 11
AAAY70126
ID AAY70126 standard; Protein; 455 AA.

XX AC AAY70126;
XX 06-JUN-2000 (first entry)
XX Human lost in leukaemia kinase (LLK).
XX Lost in leukaemia kinase; LLK; cdc2-related kinase; human; leukaemia;
KW MAPK; mitogen activated protein kinase; tumour suppressor; diagnosis;
KW chromosome 5q31.1; cytostatic; mitotic index; treatment; prevention;
KW gene therapy.
XX OS Homo sapiens.
XX Key Location/Qualifiers
FH Binding-site 11..16
FT /label= ATP_binding_site
FT Binding-site 44..50
FT /label= Cyclin_binding_motif
FT Region 125..130
FT /label= Consensus_sequence
FT /note= "Serine/threonine specific kinase sequence"
FT Domain 158..160
FT /label= MAP_kinase_activation_motif
FT Region 162..167
FT /label= Consensus_sequence
FT /note= "Serine/threonine specific kinase sequence"

XX WO200012719-A1.
XX 09-MAR-2000.
XX 31-AUG-1999; 99WO-CA00794.
XX 31-AUG-1998; 98CA-2243784.
PR 20-NOV-1998; 98CA-2251249.
XX (ONTA-) ONTARIO CANCER INST.
XX

CC leukaemia kinase (LLK alpha), isolated from rat brain cDNA library. It
CC is expressed strongly in muscle, heart, liver, brain and in tissues which
CC have a very low mitotic index. Rat LLK alpha protein has 47% and 44%
CC sequence homology to KIAMRE and KIALRE cdc2-related kinases, that are
CC putative tumour suppressors and to mitogen activated protein kinases
CC (MAPKs). This sequence has cytotostatic activity. It is useful for the
CC treatment, diagnosis and prevention of acute leukaemia and is also used
CC in gene therapy.
XX
SQ Sequence 505 AA;

Query Match 45.9%; Score 763.5; DB 21; Length 505;
Best Local Similarity 46.9%; Pred. No. 9.5e-73;
Matches 145; Conservative 65; Mismatches 92; Indels 7; Gaps 4;
QY 1 MEKYEKLAKTGEGSYGVVFKCRNKTSGQVAVVKFEVESEDDPVVKKIALREIRMLKOLKH 60
Db 1 memyetlgkvgesygvtmckkhkdtgrivaikfiye-kpekssvnkiatreikflkqfrh 59
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEPNPVGADGVIKSVLWOTLOALNFCFHIH 120
Db 60 enlvnlievfrqkkihlfvfehidhtvldelqhychgleskrlrkylfqilraieylhnn 119
QY 121 NCIHRDIKPENILITKOGIIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDYQG 179
Db 120 nihrdikpenilvsqgitklcdfgfartlaapgdvydyvatrwyrapelvldkdttyg 179
QY 180 SSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTGLKLIPRHQSFHFGISI 239
Db 180 kpvdwaiagcmiematgnpylpsssdldllhkivlkvgnltphlhnifskspifagvvl 239
QY 240 PEPEDMETLEKFSDVHPVALNFMKGCLKNPDDRLTCSQLLESSYF--DSFQE---AQI 294
Db 240 pqvqhpknarkkypknglladivhaclgdpaeirisstdllhdyftrdgmfielkfpel 299
QY 295 KRKARNEGR 303
Db 300 rakllqeak 308

RESULT 15
AAB65645
ID AAB65645 standard; Protein; 534 AA.
XX
AC AAB65645;
XX
DT 27-MAR-2001 (first entry)
XX
DE Novel protein kinase, SEQ ID NO: 172.
XX
KW Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
KW immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic;
KW dermatological; antidiabetic; antiinfertility; gene therapy; vaccine;
KW immune disorder; cardiovascular disease; neurodegenerative disease;
KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
KW inflammatory pelvic disease; multiple sclerosis; psoriasis.
XX
OS Homo sapiens.
XX
PN WO200073469-A2.
XX
PD 07-DEC-2000.
XX
PF 26-MAY-2000; 2000WO-US14842.
XX
PR 28-MAY-1999; 99US-0136503.
XX
PA (SUGE-) SUGEN INC.
XX
PI Plowman GD, Martinez R, Whyte D, Sudersanam S;
XX
DR WPI; 2001-032161/04.

DR N-PSDB; AAF44672.
XX
PT Nucleic acids encoding kinase polypeptides, useful for diagnosing and
PT treating immune-related diseases and disorders, cardiovascular disease,
PT neurodegenerative diseases and/or cancers -
XX
PS Claim 10; Fig 1; 310pp; English.
XX
CC The present sequence is a novel protein kinase. The novel protein kinases
CC and the nucleic acids that encode them may be used in the treatment and
CC diagnosis of diseases associated with inappropriate kinase expression
CC such as immune-related diseases and disorders, cardiovascular disease,
CC neurodegenerative diseases and/or cancers. The nucleic acids and
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays. The kinase polypeptides may be used as antigens in the production
CC of antibodies of kinase expression and activity. Anti-kinase antibodies
CC and kinase antagonists may also be used to down regulate kinase
CC expression and activity. Diseases related to kinase expression and
CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
CC disorders, complications of organ transplantation, myocardial infarction,
CC immune disorders, cardiomyopathies, strokes, renal failure,
CC oxidative-stress related disorders, chronic inflammatory bowel disease,
CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
CC reproductive disorders.
XX
SQ Sequence 534 AA;

Query Match 39.5%; Score 656; DB 22; Length 534;
Best Local Similarity 38.0%; Pred. No. 3.3e-61;
Matches 153; Conservative 58; Mismatches 100; Indels 92; Gaps 11;
QY 1 MEKYEKLAKTGEGSYGVVFKCRNKTSGQVAVVKFESEDDPVVKKIALREIRMLKOLKH 60
Db 1 memyetlgkvgesygvtmckkhkntgivaikfiyerpeqs-vnkiamreikfkqfhh 59
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEPNPVGADGVIKSVLWOTLOALNFCFHIH 120
Db 60 enlvnlievfrqkkihlfvfehidhtvldelqhychgleskrlrkylfqilraidylhnn 119
QY 121 N-CIHRDIKPENILITKOGIIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDYQG 178
Db 120 nviihrdikpenilvsqgitklcdfgfartlaapgdidiydyvatrwyrapelvldktsy 179
QY 179 GS--SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLI----- 214
Db 180 gkyvpvdiwalgcmiematgnpylpsssdldllhkivlkvxfmpelkakllqeakvns 239
QY 215 -----RTLGLKLIPRHQS-----IFKSNGFFHFGISI 239
Db 240 ikpkesskenelrkderktyvntllsssvlgkeiekekpkkeikvrivkvggrgdise 299
QY 240 PEPEDME---TLEEKFSVDHPVA-----LNFMKGC--LKMN-----PD 272
Db 300 pkkeyeggigggdanenvhpmspdtklvtieppninpstncnglkenphcgsgsvtmpp 359
QY 273 DRLTCSQLLESSYFDSFOEAQIKRKARNEGRNRRRQ---QVLP 312
Db 360 inltnsnlmaanlssnlfhpsvrlterakkrirtssqsigqymp 402

Search completed: May 3, 2002, 11:02:16
Job time: 51 sec

us-09-671-050-12.rag

Fri May 3 11:22:26 2002

C;Accession: S22745
R;Meyerson, M.L.
Submitted to the EMBL Data Library, May 1992
A;Reference number: S22743
A;Accession: S22745
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-376 <MEY>
A;Cross-references: EMBL:X66359
C;Genetics:
A;Introns: 152/3; 170/3
C;Superfamily: kinase-related transforming p
C;Keywords: ATP; phosphotransferase; serine/
F;3-296/Domain: protein kinase homology <KIN
F;11-19/Region: protein kinase ATP-binding mo

Query Match	70.7%;	Score 1175.5;	DB 2;	Length 376;
Best Local Similarity	67.6%;	Pred. No. 7.9e-49;		
Matches 211; Conservative	41;	Mismatches 41;		

[illegible]

RESULT 3
 S23382
 protein kinase (EC 2.7.1.37) cdk3 - human
 C;Species: Homo sapiens (man)
 C;Date: 18-Jun-1993 #sequence_revision 18-Jun-1993 #text_change 18-Jun-1999
 C;Accession: S23382; S22743
 R;Meyerson, M.; Enders, G.H.; Wu, C.L.; Su, L.K.; Gorka, C.; Nelson, E.; Tsai
 EMBO J. 11, 2909-2917, 1992
 A;Title: A family of human cdc2-related protein kinases.
 A;Reference number: S23382; MUID:92347325
 A;Accession: S23382
 A;Status: nucleic acid sequence not shown
 A;Molecule type: mRNA
 A;Residues: 1-305 <MEY>
 A;Cross-references: EMBL:X66357; NID:g36612; PIDN:CAA47001.1; PTD:g36613
 C;Genetics:
 A;Gene: GDB:CDK3
 A;Cross-references: GDB:283456
 A;Map position: 12q13-12q13
 C;Superfamily: kinase-related transforming protein; protein kinase homology
 C;Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein kinase
 F;2-25/Domain: protein kinase homology <KIN>
 F;10-18/Region: protein kinase ATP-binding motif
 F;33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 35.0%; Score 581; DB 2; Length 305;

Best Local Similarity 39.8%; Pred. No. 4.5e-21;	
Matches 123; Conservative 66; Mismatches 96; Indels 24; Gaps 8;	
QY 1 MEKYEKLAKTGE SGYGVFKCRNKTSGQWAVKKE---VESEDDPVVKKTALREIRMLKQ 57	
Db 1 MDMFQKVEKIGEGTGYGVYKAKNRETQGLVALKIRLDLEMEGVP---STAIREISLUKE 57	
QY 58 LKHPNLVNLIEVRRKRKMHLYVEFYCDHTLLNLELRNPNG-VADGVIKSVLWQTLQALNF 116	
Db 58 LKHPNIVRLDVVHNERKLYLVEFLSQDLKKYMDSTPGSELPLHLIKSYLQQLQGVSF 117	
QY 117 CHIHNCIHRDIKPENILITKQGITIKICDFGAQIL-IPGDAYTDYVATWYRAPELLVGD 175	
Db 118 CHSHRVTHRDLKPQNLLINELGAIKLADFGLARAFGVPLRTYTHEVVVTLWYRAPELLGS 177	
QY 176 TQYSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIPRHQSIFKSNQFFH 235	
Db 178 KYTTTAVDIWSIGCIFAEMVTRKALFPGDSEIDQLFRIFRMLGT-----PSEDTPW 238	
QY 236 GIS-IPE----PE-DMETLEEKFSDVHPVALNFMKGLKMNPPDRITCSQLLESSYFDS 288	
Db 229 GVTQLPDYKGSFPKWRKGLIEIVPNLEFEGRDLMLQLLOYDPSQRITAKTALAHPIYFSS 288	
QY 289 FQEAQIKRK 297	
Db 289 PEPSAARQ 297	

RESULT 4
 S40021
 protein kinase (EC 2.7.1.37) cdc2 homolog - slime mold (Dictyostelium discoideum)
 C:Species: Dictyostelium discoideum
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 18-Jun-1999
 C:Accession: S40021; S39071
 R:Michaelis, C.E.; Weeks, G.
 submitted to the EMBL Data Library, August 1992
 A:Description: The unicellular organism Dictyostelium discoideum possesses a highly
 A:Reference number: S40021
 A:Accession: S40021
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-292 <MIC>
 A:Cross-references: EMBL:L00652; NID:gl67695; PIDN:AAA16056.1; PID:gl67696
 R:Michaelis, C.; Weeks, G.
 Biochim. Biophys. Acta 1179, 117-124, 1993
 A:Title: The isolation from a unicellular organism, Dictyostelium discoideum, of a
 A:Reference number: S39071; MUID:94032415
 A:Accession: S39071
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-18,'Y',20-188,'G',190-292 <MI2>
 A:Cross-references: EMBL:L00652
 C:Complex: In various organisms, cdc2 has been identified as a component of the M-p
 C:Superfamily: Kinase-related transforming protein; protein kinase homology
 C:Keywords: ATP; cell cycle control; phosphoprotein; phosphotransferase; serine/thr
 F:2-254/Domain: protein kinase homology <KIN>
 F:10-18/Region: protein kinase ATP-binding motif
 F:33,51,126,128/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match	34.6%;	Score 574.5;	DB 2;	Length 292;
Best Local Similarity	37.5%;	Pred. No. 8.8e-21;		
Matches 115;	Conservative 64;	Mismatches 93;	Indels 35;	Gaps 5;

Qy	1	MEKYEKLAKTGESGYGVFKCRNKTSGQVAVVKFEVSEDDPVVKKIALREIRMLKQLKH	60
		: : : : : : : : :	
Db	1	MEKYSKIEKLGEGTYGIVNKAKNRETGEIVALKRIRLDSEDEGVPCTAIREISLLKELKH	60
		: : : : : : : : :	
Qy	61	PNLVNLLIEVFRKRKMHLVFEYCDHTLLNELEARNPNGVADGVIKSVLWQTILQALNFCIH	120
		: : : : : : : : :	
Db	61	PNIVRLHDVITHTERKLTLVFEYLDQDLKKYLDCEGGEISKPTIKSFMYQLLKGVAFCCHD	120
		: : : : : : : : :	
Qy	121	NCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDVAVTRWYRAPELLVGDITQVG	179
		: : : : : : : : :	

Db 1 MDDYLLKIEKIGEGTYGVVYKGRNKTTCQVAMKKIRLESEEG-VFSTAVREISLLKELQ 59

QY 60 HPNLVNLIEVFRKRKMHLVFEYCDHTLLNELERNPN--VADGVKSVLWQTLQALNFC 117

Db 60 HPNVVRLDVLMOESKLYLVFEFLSMDLKKYLDIPSQGFMDPMLVKSYLYQILEGILFC 119

QY 118 HIHNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDT 176

Db 120 HCRRVLHRDLKPQNLLIDNRGVIKLADFGLARAFGVPVRVYTHEVVTLMWYRAPEVLLGAS 179

QY 177 QYGSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLG-----KLIPRHQS 226

Db 180 RYSTPVDVWSIGTIFAELATYKPLFLHGDSEIDQLFRIFRTLGTPTNNEVWPDVESLPDYKN 239

QY 227 IF---KSNGFFHGISIPEPEDMETLEEFSDVHPVALNFMKGLKMPDDRLTCSQLLES 283

Db 240 TFPKWKSG-----NIASTVKNLDRKGIDLLTKMLIYDPPKRI SARQAMTH 284

QY 284 SYFDSFQEAQI 294

Db 285 PYFDDLDKSTL 295

RESULT 8

S22440

protein kinase (EC 2.7.1.37) cdc2 homolog 1 - rice

C;Species: Oryza sativa (rice)

C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 18-Jun-1999

C;Accession: S22440

R;Hashimoto, J.; Hirabayashi, T.; Hayano, Y.; Hata, S.; Ohashi, Y.; Suzuka, I.; Utsugi, M. Gen. Genet. 233, 10-16, 1992

A;Title: Isolation and characterization of cDNA clones encoding cdc2 homologues from Oryza sativa

A;Reference number: S22440; MUID:92293101

A;Accession: S22440

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-294 <HAS>

A;Cross-references: EMBL:X60374; NID:g20342; PIDN:CAA42922.1; PID:g20343

C;Superfamily: kinase-related transforming protein; protein kinase homology

C;Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein kinase

F;2-256/Domain: protein kinase homology <KIN>

F;10-18/Region: protein kinase ATP-binding motif

F;33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 34.2%; Score 568.5; DB 2; Length 294;

Best Local Similarity 37.6%; Pred. No. 1.7e-20;

Matches 117; Conservative 67; Mismatches 92; Indels 35; Gaps 9;

QY 1 MEKYEKLAKTGESYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60

Db 1 MEQYEKEEKIGEGTYGVVYRARDKVTNETIALKKIRLEQEDGVPSTAIRISLIKEMHH 60

QY 61 PNLVNLIEVFRKRKMHLVFEYCDHTLLN-----ELERNPNGVADGVKSVLWQTLQAL 114

Db 61 GNIVRLHDVIHSEKRIYLVFEYLDLDLKKFMDSCSEFAKNPT----LKSYLYQLRGV 115

QY 115 NFCHIHNCIHRDIKPENILITKQ-GIIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELL 172

Db 116 AYCHSHRVLHRDLKPQNLLIDRRTNALKLADFGLARAFGIPVTFTEHEVVTLMWYRAPEIL 175

QY 173 VGDYQYSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIPRHQSIFKSNG 232

Db 176 LGSROYSTPVDMSVGCIFAEMVYNQKPLFPFGDSEIDELFKIPRVLG--TPNEQS----- 227

QY 233 FEHGI-STP-----EPDMETLEEFSDVHPVALNFMKGLKMPDDRLTCSQLLE 282

Db 228 -WPGVSSLPDYKSAPFKWQAQDLATI---VPTLDPAGLDLLSKMLRYEPNKRITARQALE 283

QY 283 SSYFDSFQEAQ 293

Db 284 HEYFKDLEMVQ 294

RESULT 9

A41227

protein kinase (EC 2.7.1.37) cdk2 - human

N;Alternate names: Egl homolog; protein kinase p34

C;Species: Homo sapiens (man)

C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 18-Jun-1999

C;Accession: A41227; S17873; S16520

R;Ninomiya-Tsujii, J.; Nomoto, S.; Yasuda, H.; Reed, S.I.; Matsumoto, K. Proc. Natl. Acad. Sci. U.S.A. 88, 9006-9010, 1991

A;Title: Cloning of a human cDNA encoding a CDC2-related kinase by complementation of

A;Reference number: A41227; MUID:92020980

A;Accession: A41227

A;Molecule type: mRNA

A;Residues: 1-298 <NIN>

A;Cross-references: GB:M68520; NID:g180177; PIDN:AAA35667.1; PID:g180178

R;Tsai, L.H.; Harlow, E.; Meyerson, M. Nature 353, 174-177, 1991

A;Title: Isolation of the human cdk2 gene that encodes the cyclin A- and adenovirus E

A;Reference number: S17873; MUID:91367262

A;Accession: S17873

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-298 <TSA>

A;Cross-references: GB:X62071; NID:g312802; PIDN:CAA43985.1; PID:g312803

R;Elledge, S.J.; Spottswood, M.R. EMBO J. 10, 2653-2659, 1991

A;Title: A new human p34 protein kinase, CDK2, identified by complementation of a cdc2-

A;Reference number: S16520; MUID:91330891

A;Accession: S16520

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-176,'S',178-298 <ELL>

A;Cross-references: EMBL:X61622; NID:g29848; PIDN:CAA43807.1; PID:g29849

C;Genetics:

A;Gene: GDB:CDK2

A;Cross-references: GDB:128984; OMIM:116953

A;Map position: 12q13-12q13

C;Superfamily: kinase-related transforming protein; protein kinase homology

C;Keywords: ATP; cell cycle control; mitosis; phosphoprotein; phosphotransferase; serine/threonine-specific protein kinase

F;2-255/Domain: protein kinase homology <KIN>

F;10-18/Region: protein kinase ATP-binding motif

F;14,160/Binding site: phosphate (Thr) (covalent) #status predicted

F;15/Binding site: phosphate (Tyr) (covalent) #status predicted

F;33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 34.1%; Score 567; DB 2; Length 298;

Best Local Similarity 40.0%; Pred. No. 2e-20;

Matches 118; Conservative 60; Mismatches 99; Indels 18; Gaps 6;

QY 1 MEKYEKLAKTGESYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60

Db 1 MENFQKVEKIGEGTYGVVYKARKNLTGEVVVALKKIRLDTEGVPSTAIRISLLKELNH 60

QY 61 PNLVNLIEVFRKRKMHLVFEYCDHTLLNELERNP-NGVADGVKSVLWQTLQALNFCI 119

Db 61 PNIVKLLDVIHTENKLYLVFEFLHQDLKKFMDASALTGIPLPLTKSYLFLQLLQGLAFCHS 120

QY 120 HNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDTQY 178

Db 121 HRVLHRDLKPQNLLINTEGAIKLADFGLARAFGVPVNTYTHEVVTLMWYRAPEILLGCKYY 180

QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIPRHQSIFKSNGFFHGI- 237

Db 181 STAVDIWSLGCIFAEMVTRRALFPGDSEIDQLFRIFRTLGT-----PDEVVWPGVT 231

QY 238 STPE--PEDMETLEEFSDVHPV---ALNFMKGLKMPDDRLTCSQLLESSYF 286

Db 232 SMPDYKPSFPKWARQDFSKVVPPLDEDEGRSILSQLMLHYDPNKRISAKAALAHPPFF 286

RESULT 10

Fri May 3 11:22:28 2002

Db 181 YSTPVDIWSVGCIFAEMISQKPLFPQDSEIDQLKIFRIMGTPEYEDTWRGVTSLSLDYKSA 240

QY 228 FKSNGFFHGHSIPEDMETLEEKFSVHPVALNFMKGCLKMNPDDRLTCSQLLESSYF 286

Db 241 FPK-----WKPTDLETF---VPLNDPDGVDLLSKMLLMDPTKRINARAALAEHEYF 287

RESULT 15

A29539

protein kinase (EC 2.7.1.37) cdc2 - human

N:Alternate names: cell division control protein 2 (CDC2)

C:Species: Homo sapiens (man)

C>Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 21-Jul-2000

C:Accession: A29539

R:Lee, M.G.; Nurse, P.

Nature 327, 31-35, 1987

A:Title: Complementation used to clone a human homologue of the fission yeast cell cycle

A:Reference number: A29539; MUID:87201915

A:Accession: A29539

A:Molecule type: mRNA

A:Residues: 1-297 <LEE>

A:Cross-references: GB:X05360; NID:g29838; PIDN:CAA28963.1; PID:g29839

C:Genetics:

A:Gene: GDB:CDC2

A:Cross-references: GDB:119052; OMIM:116940

A:Map position: 10q21.1-10q21.1

C:Superfamily: kinase-related transforming protein; protein kinase homology

C:Keywords: ATP; cell cycle control; mitosis; phosphoprotein; phosphotransferase; serine

F;2-256/Domain: protein kinase homology <KIN>

F;10-18/Region: protein kinase ATP-binding motif

F;14,161/Binding site: phosphate (Thr) (covalent) #status predicted

F;15/Binding site: phosphate (Tyr) (covalent) #status predicted

F;33,51,128,130/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 33.5%; Score 556; DB 2; Length 297;

Best Local Similarity 39.4%; Pred. No. 6.6e-20;

Matches 123; Conservative 61; Mismatches 96; Indels 32; Gaps 8;

QY 1 MEKYEKLAKTGEYSYGVWFKCRNKTSGQVAVKKF-VESEDDPVVKKIALREIRMLKQLK 59

Db 1 MEDYTKIEKIGEGTYGVYKGRHKTGTQVVAMKKIRLESEEG-VPSTAIRESLLKELR 59

QY 60 HPNLVNLIEVFRKRKMHLVFEYCDHTLLNELRPNP--VADGVIKSVLMQTLQALNFC 117

Db 60 HPNIVSLQDVLMDSRLLYLIFEFLSMDLKKYLDISIPPGQYMDSSLVKSILYQILQIVFC 119

QY 118 HTHNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGD 176

Db 120 HSRRLHRDLKPQNLLDDKGTIKLADFGLARAFGIPIRVYTHEVVTWYRSPEVLLGSA 179

QY 177 QYGSSVDIWAIGCVFAELLTGQPLWPKSDVDQLYLIIRTGLKLIPIRHOSIFKSNNGFFHG 236

Db 180 RYSTPVDIWSIGTIFAELATKKPLFHGDSEIDQLFRIFRALG--TPNNE----- 226

QY 237 ISIPEPEMETLEEKFSVHPVA-----LNFMKGCLKMNPDDRLTCSQLLESS 284

Db 227 -VWPEVESLQDYKNTFFPKWPGSLASHVKNLDEGLDLSKMLIYDPAKRISGKMALNHP 285

QY 285 YFDSFQEAQIKR 296

Db 286 YFNDL-DNQIKK 296

Search completed: May 3, 2002, 11:03:17

Job time: 111 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run On: May 3, 2002, 11:08:34 ; Search time 17.01 Seconds
(without alignments)
678.978 Million cell updates/sec

Title: US-09-671-050-12
Perfect score: 1662
Sequence: 1 MEKYEKLAKTGESYGVVEK.....RKARNEGRNRRRQQLPLKS 315

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1183.5	71.2	358	1	KKIA_HUMAN
2	720.5	43.4	1030	1	STK9_HUMAN
3	581	35.0	305	1	CDK3_HUMAN
4	574.5	34.6	292	1	CC2H_DICDI
5	571.5	34.4	297	1	CDK2_XENLA
6	570.5	34.3	302	1	CC2_CARAU
7	568.5	34.2	294	1	CC21_ORYSA
8	567	34.1	298	1	CDK2_HUMAN
9	566.5	34.1	294	1	CC2_MAIZE
10	563	33.9	298	1	CDK2_CRIGR
11	561.5	33.8	294	1	CC2A_ARATH
12	561	33.8	298	1	CDK2_RAT
13	559	33.6	298	1	CDK2_MESAU
14	559	33.6	302	1	CC21_XENLA
15	557	33.5	297	1	CC2_RAT
16	556	33.5	297	1	CC2_BOVIN
17	556	33.5	297	1	CC2_HUMAN
18	556	33.5	302	1	CC22_XENLA
19	554	33.3	301	1	CC21_TRYBB
20	553	33.3	298	1	CDK2_CARAU
21	551.5	33.2	294	1	CC2_VIGUN
22	551.5	33.2	311	1	CC23_TRYBB
23	551	33.2	288	1	CC2H_PLAFK
24	548	33.0	297	1	CC2_MOUSE
25	547.5	32.9	303	1	CC2_CHICK
26	546.5	32.9	297	1	CC2_DROME
27	545	32.8	292	1	CDK5_BOVIN
28	545	32.8	292	1	CDK5_RAT
29	544	32.7	292	1	CDK5_HUMAN
30	543	32.7	292	1	CDK5_MOUSE
31	543	32.7	301	1	CC21_TRYCO
32	541.5	32.6	294	1	CC22_MEDSA
33	540	32.5	296	1	CC2_DICDI

34	536	32.3	292	1	CDK5_XENLA
35	535.5	32.2	294	1	CC2A_ANTMA
36	533	32.1	305	1	PH85_YEAST
37	530.5	31.9	346	1	CDK7_HUMAN
38	529	31.8	294	1	CDK5_DROME
39	529	31.8	346	1	CDK2_MOUSE
40	528	31.8	292	1	CC22_ORYSA
41	527.5	31.7	346	1	CDK7_MOUSE
42	525	31.6	332	1	CC2_CAEEL
43	519.5	31.3	314	1	CC2C_DROME
44	519	31.2	345	1	CC22_TRYBB
45	518.5	31.2	297	1	CC2_SCHPO

P51166	xenopus lae
Q38772	antirrhinum
P17157	saccharomyc
P50613	homo sapien
P48609	drosophila
P97377	mus musculus
P29619	oryza sativ
Q03147	mus musculus
P34556	caenorhabdi
P23573	drosophila
P54665	trypanosoma
P04551	schizosacch

ALIGNMENTS

RESULT 1	
KKIA_HUMAN	
ID	KKIA_HUMAN
AC	Q00532;
DT	01-APR-1993 (Rel. 25, Created)
DT	01-APR-1993 (Rel. 25, Last sequence update)
DT	20-AUG-2001 (Rel. 40, Last annotation update)
DE	SERINE/THREONINE-PROTEIN KINASE KIALRE (EC 2.7.1.-) (CYCLIN-DEPENDENT KINASE-LIKE 1).
DE	KINASE-LIKE 1).
GN	CDKL1.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=92347325; PubMed=1639063;
RA	Meyerson M., Enders G.H., Wu C.-L., Su L.-K., Gorka C., Nelson C.,
RA	Harlow E., Tsai L.-H.;
RT	"A family of human cdc2-related protein kinases.";
RL	EMBO J. 11:2909-2917(1992).
CC	-!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC	CDC2/CDKX SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X66358; CAA47002.1; .	
EMBL; X66359; CAA47002.1; JOINED.	
PIR; S22744; S22744.	
PIR; S22745; S22745.	
PIR; S23383; S23383.	
HSSP; P24941; 1AQL.	
MTM; 603441; .	
InterPro; IPR000719; Euk_pkinase.	
InterPro; IPR002290; Ser_thr_kin_actsite.	
Pfam; PF00069; pkinase; 1.	
SMART; SM00220; S_TKc; 1.	
PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.	
PROSITE; PS00108; PROTEIN_KINASE_ST; 1.	
PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.	
Transferase; Serine/threonine-protein kinase; ATP-binding.	
DOMAIN 5 288	PROTEIN KINASE.
NP_BIND 11 19	ATP (BY SIMILARITY).
BINDING 34 34	ATP (BY SIMILARITY).
ACT_SITE 127 127	BY SIMILARITY.
SEQUENCE 358 AA; 41834 MW; 88344321F24B77C6 CRC64;	

Query Match 71.2%; Score 1183.5; DB 1; Length 358;
Best Local Similarity 71.8%; Pred. No. 1.le-78;

Matches 211; Conservative 41; Mismatches 41; Indels 1; Gaps 1;

QY 1 MEKYEKLAKTGEGSYGVVFKCRNKTSGQVVAVKKFVSEDDPVVKKIALREIRMLKQLKH 60
Db 2 MEKYEKIGKIGEGSYGVVFKCRNRDQTQVIAIKFESEDDPVIKKIALREIRMLKQLKH 61

QY 61 PNLVNLIEVFRRRKRMHLVFEYCDHTLLELERNPNGVADGVIKSVLWQTLQALNFCCHI 120
Db 62 PNLVNLIEVFRRRKRLHLVFEYCDHTVLHLDLRYQGVPEHLVKLSITWQTLQAVNFCCHK 121

QY 121 NCIHRIKIPENILITKQGIKICDFGFAQILI-PGDAYTDYVATRWYRAPPELLVGDTOYG 179
Db 122 NCIHRIKIPENILITKHSVIKLCDFGFARLLTGPSDYTDYVATRWYRSPPELLVGDTOYG 181

QY 180 SSVDIWAIGCVFAELLTGQPLWPGKSDVDQYLIIRTLGKLIPRHQSIFKSNGFFHGISI 239
Db 182 PPVDVWAIGCVFAELLSGVPLWPGKSDVDQYLIIRKTLGDLIPRHQQVFNSTNQYFSGVKI 241

QY 240 PEPEDMETLEKFSVDVHPVALNFMKGCLKMNPDRLTCSQLLESSYFDSFQEAQ 293
Db 242 PDPEDMEPLKLPFNISYPALGLLKGCLHMDPRLTCEQLLHPYFENIREIE 295

RESULT 2
STK9_HUMAN STANDARD; PRT; 1030 AA.
AC O76039; Q14198;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SERINE/THREONINE-PROTEIN KINASE 9 (EC 2.7.1.37).
GN STK9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98389628; PubMed=9721213;
RA Montini E., Andolfi G., Caruso A., Buchner G., Walpole S.M.,
RA Mariani M., Consalez G.G., Trump D., Ballabio A., Franco B.;
RT "Identification and characterization of a novel serine-threonine
RT kinase gene from the xp22 region.";
RL Genomics 51:427-433(1998).
RN [2]
RP SEQUENCE OF 339-789 FROM N.A.
RA Krause S.W., Rehli M., Kreutz M., Schwarzfischer L., Paulauskis J.D.,
RA Andreessen J.D.;
RT "Differential screening leads to novel genetic markers of monocyte
RT to macrophage maturation.";
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN = ADP + A PHOSPHOPROTEIN.
CC -!- ENZYME REGULATION: COULD BE ACTIVATED BY PHOSPHORYLATION ON
CC TYROSINE AND THREONINE. PROBABLY BOTH PHOSPHORYLATIONS ARE
CC REQUIRED FOR ACTIVITY.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, LUNG, KIDNEY, PROSTATE,
CC OVARY, PLACENTA, PANCREAS AND TESTIS.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC -!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-10 IS THE INITIATOR.
CC -!- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 415.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; Y15057; CAA75342.1; -.
DR EMBL; X89059; CAA61445.1; ALT_FRAME.

MIM: 300203; -.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_kin_actsite.
DR Pfam: PF00069; pkinase; 1.
DR SMART: SMO0220; S_TKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW transferase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation.
FT DOMAIN 13 297 PROTEIN KINASE.
FT NP_BIND 19 27 ATP (BY SIMILARITY).
FT BINDING 42 42 ATP (BY SIMILARITY).
FT ACT_SITE 135 135 BY SIMILARITY.
FT MOD_RES 169 169 PHOSPHORYLATION (ACTIVATES THE KINASE)
FT (POTENTIAL).
FT MOD_RES 171 171 PHOSPHORYLATION (ACTIVATES THE KINASE)
FT (POTENTIAL).
FT DOMAIN 784 789 POLY-LYS.
FT CONFLICT 339 340 HR -> GT (IN REF. 2).
FT CONFLICT 541 541 L -> W (IN REF. 2).
FT CONFLICT 731 764 MISSING (IN REF. 2).
SQ SEQUENCE 1030 AA; 115537 MW; 8A1C9C438610EF08 CRC64;

Query Match 43.4%; Score 720.5; DB 1; Length 1030;
Best Local Similarity 44.4%; Pred. No. 9.7e-45;
Matches 139; Conservative 69; Mismatches 98; Indels 7; Gaps 4;

QY 1 MEKYEKLAKTGEGSYGVVFKCRNKTSGQVVAVKKFVSEDDPVVKKIALREIRMLKQLKH 60
Db 10 MNKFEILGVVGAGYGVVLCRHKETHETHEIVAIAKKFKDSEENEVEKETTTLRELKMLRTLKQ 69

QY 61 PNLVNLIEVFRRRKRMHLVFEYCDHTLLELERNPNGVADGVIKSVLWQTLQALNFCCHI 120
Db 70 ENIVELKEAFRRRGKLYLVFEYVEKNNLLEEMPNGVPEKVKSYIYQLIAIHWCKN 129

QY 121 NCIHRIKIPENILITKQGIKICDFGFAQILIPGD--AYTDYVATRWYRAPPELLVGDTOY 178
Db 130 DIVHRDIKIPENILISHNDVLKLCDFGFARNLSEGNANYTEYVATRWYRSPPELLLG-APY 188

QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQYLIIRTLGKLIPRHQSIFKSNGFFHGIS 238
Db 189 GKSVDMWVSGCILGELSQQPLFPGESEIDQLFTTQKVLGPLPSEQMKLFYSNPRFHGLR 248

QY 239 IPEPEMETLEEKFSVDV-HPVALNFMKGCLKMNPDRLTCSQLLESSYFDSFQEAQIKRK 297
Db 249 FPAVNHPQSLERRYLGIILNSVLLDLKMLLLKLPADRYLTEQCLN---HPTFTQTRLDDR 305

QY 298 ARNEGNRRRQQV 310
Db 306 SPFSRAKRPYHV 318

RESULT 3
CDK3_HUMAN STANDARD; PRT; 305 AA.
ID CDK3_HUMAN STANDARD; PRT; 305 AA.
AC Q00526;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CELL DIVISION PROTEIN KINASE 3 (EC 2.7.1.-).
GN CDK3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=92347325; PubMed=1639063;
RA Meyerson M., Enders G.H., Wu C.-L., Su L.-K., Gorka C., Nelson C.,
RA Harlow E., Tsai L.-H.;


```
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D17758; BAA04605.1; -.
CC HSSP; P24941; LHCK.
CC DR InterPro; IPR000719; Euk_pkinase.
CC DR InterPro; IPR002290; Ser_thr_kin_actsite.
CC DR pfam; PF00069; pkinase; 1.
CC DR SMART; SM00220; S_TKc; 1.
CC DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC DR Transferrase; Serine/threonine-protein kinase; ATP-binding;
KW Cell cycle; Cell division; Mitosis; Nuclear protein; Phosphorylation.
FT DOMAIN 4 287 PROTEIN KINASE.
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 128 128 BY SIMILARITY.
FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 161 161 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).
SQ SEQUENCE 302 AA; 34499 MW; 58DB812E19B311F5 CRC64;

Query Match 34.3%; Score 570.5; DB 1; Length 302;
Best Local Similarity 38.6%; Pred. No. 1.7e-34;
Matches 120; Conservative 68; Mismatches 90; Indels 33; Gaps 7;

QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVAVKKF-VESEDDPVVKKIALREIRMLKQLK 59
Db 1 MDDYLKIEKIGEGTYGVVYKGNKTTGQVAVAMKKIRLESEEG-VPSTAVREISLLKELQ 59

QY 60 HPNLVNLIEVFRKRKMHVFEYCDHTLLNELERNPNG--VADGVKSVLWQTLQALNFC 117
Db 60 HPNVRLDLVLMQESKLYLVFEFLSMDLKKYLDIPSQGFMDPMLVKSILYQILEGILFC 119

QY 118 HHNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGD 176
Db 120 HCRVLRDLKPQNLLIDNKGVIKLADEGLARAFGVPVRVYTHEVVTLWYRAPEVLLGAS 179

QY 177 QYGSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLG-----KLIPRHQS 226
Db 180 RYSTPVDVWSIGTIFAELATKPLPLFGDSEIDQLFRIFRTLGTPTNNEVPDVESLPDYKN 239

QY 227 IF--KSNGFFHGISIPEPEMETLEEKFSVHPVVALNFMKGCLKMNPDRLTCSOLLES 283
Db 240 TFPKWKSG-----NLASTVKNLDKNGIDLLTKMLIYDPPKRISARQAMTH 284

QY 284 SYFDSFQEAQI 294
Db 285 PYFDDLKSTL 295

RESULT 7
CC21_ORYSA
ID CC21_ORYSA STANDARD; PRT; 294 AA.
AC P29618;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE CELL DIVISION CONTROL PROTEIN 2 HOMOLOG 1 (EC 2.7.1.-).
GN CDC2-1.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RX MEDLINE=92293101; PubMed=1376401;
```

```
RA Hashimoto J., Hirabayashi T., Hayano Y., Hata S., Ohashi Y.,
RA Suzuki I., Utsugi T., Toh-E A., Kikuchi Y.,
RT "Isolation and characterization of cDNA clones encoding cdc2
RT homologues from Oryza sativa: a functional homologue and cognate
RT variants.";
RL Mol. Gen. Genet. 233:10-16(1992).
CC -!- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC
CC CELL CYCLE. COMPONENT OF THE KINASE COMPLEX THAT PHOSPHORYLATES
CC THE REPETITIVE CARBOXYL-TERMINUS OF RNA POLYMERASE II.
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-161 ACTIVATES IT (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X60374; CAA42922.1; -.
CC PIR; S22440; S22440.
CC HSSP; P24941; LHCK.
CC DR InterPro; IPR000719; Euk_pkinase.
CC DR InterPro; IPR002290; Ser_thr_kin_actsite.
CC DR pfam; PF00069; pkinase; 1.
CC DR SMART; SM00220; S_TKc; 1.
CC DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC Transferrase; Serine/threonine-protein kinase; ATP-binding;
KW Cell cycle; Cell division; Mitosis; Phosphorylation.
FT DOMAIN 4 287 PROTEIN KINASE.
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 127 127 BY SIMILARITY.
FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 161 161 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).
SQ SEQUENCE 294 AA; 34071 MW; 51322D93AEF4C131 CRC64;

Query Match 34.2%; Score 568.5; DB 1; Length 294;
Best Local Similarity 37.6%; Pred. No. 2.3e-34;
Matches 117; Conservative 67; Mismatches 92; Indels 35; Gaps 9;

QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEQYEKEEKIGEGTYGVVYRARDKVTNETIALKTRLEQDEGVPSTAIRISLLKEMHH 60

QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLN-----ELERNPNGVADGVKSVLWQTLQAL 114
Db 61 GNIIVLRHDVHSEKRIYLVFEYLDLDLKKFMDSCPFAKNPT-----LIKSYLYQILRGV 115

QY 115 NFCHTHNCIHRDIKPENILITKQ-GIIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELL 172
Db 116 AYCHSHRVLHRDLKPQNLLIDRRTNALKLADFGLARAFGIPVTRTHEVVTLWYRAPEIL 175

QY 173 VGDYQYSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIPRHQSIFKSNG 232
Db 176 LGSRQYSTPVDMMWSVGCIFAEMVNQKPLFPDGSEIDELEFKIFRVLG--TPNEQS----- 227

QY 233 FFHGI-SIP-----EPEDMETLEEKFSDVHPVVALNFMKGCLKMNPDRLTCSOLLE 282
Db 228 -WPGVSSLPDYKSAFPKWQAQDLATI---VPTLDPAGLDLLSKMLRYEPNKRITARQALE 283

QY 283 SSYFDSFQEAQ 293
Db 284 HEYFKDLEMVQ 294
```

RESULT 8
CDK2_HUMAN
ID CDK2_HUMAN STANDARD; PRT; 298 AA.
AC P24941;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CELL DIVISION PROTEIN KINASE 2 (EC 2.7.1.-) (P33 PROTEIN KINASE).
GN CDK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91330891; PubMed=1714386;
RA Elledge S.J., Spottswood M.R.;
RT "A new human p34 protein kinase, CDK2, identified by complementation
of a cdc28 mutation in Saccharomyces cerevisiae, is a homolog of
Xenopus Egl.";
RL EMBO J. 10:2653-2659(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91367262; PubMed=1653904;
RA Tsai L.-H., Harlow E., Meyerson M.;
RT "Isolation of the human cdk2 gene that encodes the cyclin A- and
adenovirus E1A-associated p33 kinase.";
RL Nature 353:174-177(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92020980; PubMed=1717994;
RA Ninomiya-Tsuji J., Nomoto S., Yasuda H., Reed S.I., Matsumoto K.;
RT "Cloning of a human cDNA encoding a CDC2-related kinase by
complementation of a budding yeast cdc28 mutation.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:9006-9010(1991).
RN [4]
RP PHOSPHORYLATION SITES.
RX MEDLINE=93010995; PubMed=1396589;
RA Gu Y., Rosenblatt J., O'Morgan D.O.;
RT "Cell cycle regulation of CDK2 activity by phosphorylation of Thr160
and Tyr15.";
RL EMBO J. 11:3995-4005(1992).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RX MEDLINE=93288132; PubMed=8510751;
RA de Bondt H.L., Rosenblatt J., Jancarik J., Jones H.D.,
RA Morgan D.O., Kim S.-H.;
RT "Crystal structure of cyclin-dependent kinase 2.";
RL Nature 363:595-602(1993).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF COMPLEX WITH CYCLIN A.
RX MEDLINE=95356811; PubMed=7630397;
RA Jeffrey P.D., Russo A.A., Polyak K., Gibbs E., Hurwitz J.,
RA Massague J., Pavletich N.P.;
RT "Mechanism of CDK activation revealed by the structure of a
cyclinA-CDK2 complex.";
RL Nature 376:313-320(1995).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.33 ANGSTROMS) OF COMPLEX WITH L868276.
RX MEDLINE=96181476; PubMed=8610110;
RA de Azevedo W.F. Jr., Muleer-Dieckmann H.-J., Schulze-Gahmen U.,
RA Worland P.J., Sausville E., Kim S.-H.;
RT "Structural basis for specificity and potency of a flavonoid
inhibitor of human CDK2, a cell cycle kinase.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:2735-2740(1996).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF COMPLEX WITH CG2A AND KIP1.
RX MEDLINE=96300318; PubMed=8684460;
RA Russo A.A., Jeffrey P.D., Patten A.K., Massague J., Pavletich N.P.;
RT "Crystal structure of the p27Kip1 cyclin-dependent-kinase inhibitor
bound to the cyclin A-Cdk2 complex.";
RL Nature 382:325-331(1996).

[9]
RN X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF COMPLEX WITH CG2A.
RX MEDLINE=96313126; PubMed=8756328;
RA Russo A.A., Jeffrey P.D., Pavletich N.P.;
RT "Structural basis of cyclin-dependent kinase activation by
phosphorylation.";
RL Nat. Struct. Biol. 3:696-700(1996).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=97075215; PubMed=8917641;
RA Schulze-Gahmen U., de Bondt H.L., Kim S.-H.;
RT "High-resolution crystal structures of human cyclin-dependent kinase
2 with and without ATP: bound waters and natural ligand as guides for
inhibitor design.";
RL J. Med. Chem. 39:4540-4546(1996).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=97475219; PubMed=9334743;
RA Lawrie A.M., Noble M.E.M., Tunnah P., Brown N.R., Johnson L.N.,
RA Endicott J.A.;
RT "Protein kinase inhibition by staurosporine revealed in details of
the molecular interaction with CDK2.";
RL Nat. Struct. Biol. 4:796-801(1997).
RN [12]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF COMPLEX WITH CKS1.
RX MEDLINE=96182647; PubMed=8601310;
RA Bourne Y., Watson M.H., Hickey M.J., Holmes W., Rocque W., Reed S.I.,
RA Tainer J.A.;
RT "Crystal structure and mutational analysis of the human CDK2 kinase
complex with cell cycle-regulatory protein CksHs1.";
RL Cell 84:863-874(1996).
RN [13]
RP X-RAY CRYSTALLOGRAPHY (2.05 ANGSTROMS).
RX MEDLINE=98342369; PubMed=9677190;
RA Gray N.S., Wodicka L., Thunnissen A.-M.W.H., Norman T.C., Kwon S.,
RA Espinoza F.H., Morgan D.O., Barnes G., Leclerc S., Meijer L.,
RA Kim S.H., Lockhart D.J., Schultz P.G.;
RT "Exploiting chemical libraries, structure, and genomics in the search
for kinase inhibitors.";
RL Science 281:533-538(1998).
CC -!- FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.
CC INTERACTS WITH CYCLINS A, D, OR E. ACTIVITY OF CDK2 IS MAXIMAL
CC DURING S PHASE AND G2.
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-160 ACTIVATES IT.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X61622; CAA43807.1; -
DR EMBL; X62071; CAA43985.1; -
DR EMBL; M68520; AAA35667.1; -
DR PIR; A41227; A41227.
DR PIR; S16520; S16520.
DR PIR; S17873; S17873.
DR PDB; 1FIN; 27-JAN-97.
DR PDB; 1HCK; 07-DEC-96.
DR PDB; 1HCL; 07-DEC-96.
DR PDB; 1AQ1; 12-NOV-97.
DR PDB; 1JST; 11-JAN-97.
DR PDB; 1JSU; 29-JUL-97.
DR PDB; 1BUH; 09-SEP-98.
DR PDB; 1B38; 23-DEC-98.
DR PDB; 1B39; 23-DEC-98.
DR PDB; 1CKP; 13-JAN-99.
DR MIM; 116953; -.

```
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_kin_actsite.
DR Pfam; PF00069; pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Cell cycle; Cell division; Mitosis; Phosphorylation; 3D-structure.
FT DOMAIN 4 286
FT NP_BIND 10 18
FT BINDING 33 33
FT ACT_SITE 127 127
FT MOD_RES 14 14
FT MOD_RES 15 15
FT MUTAGEN 14 14
FT MUTAGEN 15 15
FT MUTAGEN 160 160
SQ SEQUENCE 298 AA; 33929 MW; F90A0F4E70910B51 CRC64;

Query Match 34.1%; Score 567; DB 1; Length 298;
Best Local Similarity 40.0%; Pred. No. 3e-34;
Matches 118; Conservative 60; Mismatches 99; Indels 18; Gaps 6;

QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MENFORKEKIGEGTYGVVYKARNKLTGEVVALKKIRLDTETEGVPSTAIRISLLKELNH 60

QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNLELRNP-NGVADGVIKSVLWQTLQALNFC 119
Db 61 PNIVKLLDVHTENKLYLVFEFLHQDLKKFMDASALTGIPLPKISYLFQLQGLAFCHS 120

QY 120 HNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELIVGDTQY 178
Db 121 HRVLRDLKPQNLLINTEGAIKLADEGLARAFGVPVRTYTHEVVTWYRAPEILLGCKYY 180

QY 179 GSSVDIWAIGCVFAELLTGQPLWPKSDVDQLYLIIRTLGKLIIPRHQSIFKSNFFHGI- 237
Db 181 STAVDIWSLGCIFAEMVTRRALFPDSEIDLQLFRIFRITLGT-----PDEVVWPGVT 231

QY 238 SIPE--PEDMETLEEKFSVDVHPV----ALNFMKGCLKMNPDDRDLTCSQLLESSYF 286
Db 232 SMPDYKPSFPKWARQDFSKVVPPLDDEGRSLLSQMLHYDPNKRISAKAALAHPPF 286

RESULT 9
CC2_MAIZE STANDARD; PRT; 294 AA.
AC P23111;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CELL DIVISION CONTROL PROTEIN 2 HOMOLOG (EC 2.7.1.-).
GN CDC2.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoidae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91195354; PubMed=2014258;
RA Colasanti J., Tyers M., Sundaresan V.;
RT "Isolation and characterization of cDNA clones encoding a functional
RT p34cdc2 homologue from Zea mays.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:3377-3381(1991).
CC -!- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC
CC CELL CYCLE. COMPONENT OF THE KINASE COMPLEX THAT PHOSPHORYLATES
CC THE REPEITIVE CARBOXYL-TERMINUS OF RNA POLYMERASE II.
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-161 ACTIVATES IT (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M60526; AAA33479.1; -.
CC HSSP; P24941; 1HCK.
CC MaizeDB; 60686; -.
CC InterPro; IPR000719; Euk_pkinase.
CC InterPro; IPR002290; Ser_thr_kin_actsite.
CC Pfam; PF00069; pkinase; 1.
CC SMART; SM00220; S_TKC; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Cell cycle; Cell division; Mitosis; Phosphorylation.
KW DOMAIN 4 287
KW NP_BIND 10 18
KW BINDING 33 33
KW ACT_SITE 127 127
KW MOD_RES 14 14
KW MOD_RES 15 15
KW MOD_RES 161 161
SQ SEQUENCE 294 AA; 33834 MW; 5063ECFCC2D5FFDD CRC64;

Query Match 34.1%; Score 566.5; DB 1; Length 294;
Best Local Similarity 37.6%; Pred. No. 3.2e-34;
Matches 117; Conservative 66; Mismatches 93; Indels 35; Gaps 9;

QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEQYEKVEKIGEGTYGVVYKALDKATNETIALKKIRLEQDEGVPSSTAIRISLLKENNH 60

QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLN-----ELERNPNGVADGVIKSVLWQTLQAL 114
Db 61 GNIVRLHDVVHSEKRIYLVFEYLDLDLKKFMDSCFEFAKNPT-----LIKSYLYQILHGV 115

QY 115 NFCHIHNCIHRDIKPENILITKQ-GIHKICDFGFAQIL-IPGDAYTDYVATRWYRAPEL 172
Db 116 AYCHSHRVLHRDLKPQNLLIDRRTNALKLADEGLARAFGIPVRTFHEVVTWYRAPEIL 175

QY 173 VGDQYQSSVDIWAIGCVFAELLTGQPLWPKGSDVDQLYLIIRTLGKLIIPRHQSIFKSNG 232
Db 176 LGARQYSTPVDVWWSVGCIFAEMVNQKPLFPDSEIDELFKIFRLG--TPNEQS----- 227

QY 233 FFHGIS-IP-----EPEDMETLEEKFSVDVHPVALNFMKGCLKMNPDDRDLTCSQLE 282
Db 228 -WPGVSCLPDFKTAFFRWQAQDLATV---VPNLDPAGLDLLSKMLRYEPSKRITARQALE 283

QY 283 SSYFDSFQEAQ 293
Db 284 HEYFKDLEVVQ 294

RESULT 10
CDK2_CRIGR STANDARD; PRT; 298 AA.
ID CDK2_CRIGR
AC O55076;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CELL DIVISION PROTEIN KINASE 2 (EC 2.7.1.-).
GN CDK2.
OS Cricetulus griseus (Chinese hamster).
```



```
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94107363; PubMed=8280171;
RA Noguchi E., Sekiguchi T., Yamashita K., Nishimoto T.;
RT "Molecular cloning and identification of two types of hamster cyclin-
RT dependent kinases: cdk2 and cdk2L.";
RL Biochem. Biophys. Res. Commun. 197;1524-1529(1993).
CC -!- FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.
CC INTERACTS WITH CYCLINS A, D, OR E. ACTIVITY OF CDK2 IS MAXIMAL
CC DURING S PHASE AND G2.
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-160 ACTIVATES IT (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D17350; BAA04165.1; -
DR HSSP; P24941; 1AQ1.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_kin_actsite.
DR Pfam; PF00069; pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Cell cycle; Cell division; Mitosis; Phosphorylation.
FT DOMAIN 4 286
FT NP_BIND 10 18
FT BINDING 33 33
FT ACT_SITE 127 127
FT MOD_RES 14 14
FT MOD_RES 15 15
FT MOD_RES 160 160
SQ SEQUENCE 298 AA; 33839 MW; D1LC036664C4C7CE CRC64;
-----
Query Match 33.6%; Score 559; DB 1; Length 298;
Best Local Similarity 39.7%; Pred. No. 1.le-33;
Matches 117; Conservative 60; Mismatches 100; Indels 18; Gaps 6;
QY 1 MEKYEKLAKTGEGSYGVVPKCRNKTSGQWAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db ||:|:| ||:|:|:| ||:|:|:|:| ||:|:|:|:| ||:|:|:|:|:| ||:|:|:|:|:|
1 MENFQKVEKIGEGTYGVVYKAKNKLTEGEVWALKKIRLDTETEGVSPSTAIREISLLKELNH 60
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRNP-NGVADGVKSVLWQTLQALNFCCHI 119
Db ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|
61 PNIVKLLDVHTENKLYLVFELLHQDLKKFMDASAVTGIPLKSYLQQLQGLAFCHS 120
QY 120 HNCIHRDIKPENILITKQGIKICDFGFAQL-IPGDAYTDYVATRWYRAPELLVGDQY 178
Db | :||:|:|:|:|:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|
121 HRVLHRDLKPQNLLINAECSIKLADFGLARAFGVVRTYTHEVTLWYRAPELLGCKYY 180
QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIIPRHOSIFKSNFFHGI- 237
Db ::|||:|:|:|:|:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|
181 STAVDINSLGCIFAEMVTRRALFPDSEIDQLFRIFRTLGT-----PDEVVWPGVT 231
QY 238 SIPE--PEDMETLEKFSVDVHPV----ALNFMKGCLKNPDDRLTCSQLLESSYF 286
Db |:|:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|
232 SMPDYKPSFPKWARQDFSKVVPPLDEDEGRSLLSQMLHYDPNKRISAKAALAHPPF 286
```

```
RESULT 14
CC21_XENLA
ID CC21_XENLA STANDARD; PRT; 302 AA.
AC P35567;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CELL DIVISION CONTROL PROTEIN 2 HOMOLOG 1 (EC 2.7.1.1-) (P34 PROTEIN
DE KINASE).
GN CDC2X1.1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Oocyte;
RX MEDLINE=92318937; PubMed=1377775;
RA Pickham K.M., Meyer A.N., Li J., Donoghue D.J.;
RT "Requirement of mosXe protein kinase for meiotic maturation of
RT Xenopus oocytes induced by a cdc2 mutant lacking regulatory
RT phosphorylation sites.";
RL Mol. Cell. Biol. 12:3192-3203(1992).
RN [2]
RP PHOSPHORYLATION OF THR-161.
RX MEDLINE=93345456; PubMed=8344251;
RA Fesquet D., Labbe J.-C., Derancourt J., Capony J.-P., Galas S.,
RA Girard F., Lorca T., Shuttleworth J., Doree M., Cavadore J.-C.;
RT "The Mol5 gene encodes the catalytic subunit of a protein kinase that
RT activates cdc2 and other cyclin-dependent kinases (CDKs) through
RT phosphorylation of Thr161 and its homologues.";
RL EMBO J. 12:3111-3121(1993).
CC -!- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC
CC CELL CYCLE. IT IS REQUIRED IN HIGHER CELLS FOR ENTRY INTO S-PHASE
CC AND MITOSIS. COMPONENT OF THE KINASE COMPLEX THAT PHOSPHORYLATES
CC THE REPETITIVE CARBOXYL-TERMINUS OF RNA POLYMERASE II (BY
CC SIMILARITY).
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-161 ACTIVATES IT.
CC -!- SUBUNIT: FORMS A STABLE BUT NON-COVALENT COMPLEX WITH A REGULATORY
CC SUBUNIT AND WITH A CYCLIN.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M60680; AAA63561.1; -
DR PIR; A44349; A44349.
DR HSSP; P24941; LHCK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_kin_actsite.
DR Pfam; PF00069; pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Cell cycle; Cell division; Mitosis; Nuclear protein; Phosphorylation.
FT DOMAIN 4 287
FT NP_BIND 10 18
FT BINDING 33 33
FT ACT_SITE 128 128
FT MOD_RES 14 14
FT PHOSPHORYLATION (BY SIMILARITY).
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2002, 11:08:12 ; Search time 40.57 Seconds
(without alignments)
1135.711 Million cell updates/sec

Title: US-09-671-050-12
Perfect score: 1662
Sequence: 1 MEKYEKLAKTGESYGVVFK.....RKARNEGRNRRRQQVLPLKS 315

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues 473505
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_17:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1027.5	61.8	353	5 Q9U2H1	Q9u2h1 caenorhabdi
2	1027.5	61.8	392	5 Q9VMN3	Q9vmn3 drosophila
3	967.5	58.2	566	6 Q9TTK0	Q9ttk0 oryctolaqus
4	941.5	56.6	493	4 Q92772	Q92772 homo sapien
5	930.5	56.0	564	11 Q9QYI2	Q9qyi2 mus musculu
6	930.5	56.0	568	11 Q9QUK0	Q9quk0 mus musculu
7	928	55.8	329	11 Q9QYI1	Q9qyi1 mus musculu
8	784.5	47.2	455	4 Q9P114	Q9p114 homo sapien
9	781	47.0	657	5 Q9BMT9	Q9bmt9 leishmania
10	769.5	46.3	1106	5 Q9JMT0	Q9jmt0 rattus norv
11	763.5	45.9	457	11 Q9JMT2	Q9jmt2 rattus norv
12	763.5	45.9	505	11 Q9JMT1	Q9jmt1 rattus norv
13	644	38.7	1104	13 Q9W6R6	Q9w6r6 fugu rubrip
14	641.5	38.6	997	4 Q9UJL6	Q9ujl6 homo sapien
15	579.5	34.9	294	10 Q9ZRI1	Q9zri1 triticum ae
16	567.5	34.1	303	13 Q9DGA2	Q9dga2 oryzias jav
17	564.5	34.0	303	13 Q9DGA1	Q9dga1 oryzias jav
18	564	33.9	288	5 Q9XZD6	Q9xzd6 plasmodium
19	562	33.8	288	5 Q96820	Q96820 plasmodium

20	561.5	33.8	303	13 Q9DGD3	Q9dgd3 oryzias lat
21	560	33.7	288	5 Q96821	Q96821 plasmodium
22	559.5	33.7	303	13 Q9DGA5	Q9dga5 oryzias cur
23	559.5	33.7	303	13 Q9DG98	Q9dgg98 oryzias luz
24	558.5	33.6	294	10 P93101	P93101 chenopodium
25	556.5	33.5	294	10 Q9M307	Q9m307 arabidopsis
26	555.5	33.4	300	3 Q13379	Q13379 pneumocysti
27	555.5	33.4	300	3 Q13380	Q13380 pneumocysti
28	555.5	33.4	300	5 Q15890	Q15890 toxoplasma
29	554.5	33.4	294	10 Q82666	Q82666 brassica na
30	554.5	33.4	294	10 Q40790	Q40790 pinus conto
31	552.5	33.2	300	5 Q44000	Q44000 toxoplasma
32	551	33.2	292	13 Q9DE44	Q9de44 brachydanio
33	550.5	33.1	294	10 Q41639	Q41639 vigna acon
34	550.5	33.1	294	10 Q43361	Q43361 picea abies
35	549.5	33.1	294	5 Q76541	Q76541 cryptospori
36	549.5	33.1	294	10 Q40484	Q40484 nicotiana t
37	549.5	33.1	294	10 Q9FUR4	Q9fur4 nicotiana t
38	549	33.0	298	5 Q27032	Q27032 theileria p
39	547.5	32.9	294	10 P93556	P93556 sesbania ro
40	547.5	32.9	300	5 Q17066	Q17066 asterina pe
41	547	32.9	298	5 Q26671	Q26671 theileria a
42	546.5	32.9	297	5 Q9TX74	Q9tx74 drosophila
43	546.5	32.9	297	5 Q9TX69	Q9tx69 drosophila
44	545.5	32.8	297	5 Q9TX73	Q9tx73 drosophila
45	544.5	32.8	294	10 Q65838	Q65838 lycopersico

ALIGNMENTS

RESULT 1
Q9U2H1 ID Q9U2H1 PRELIMINARY; PRT; 353 AA.
AC Q9U2H1;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE Y42A5A.4 PROTEIN.
GN Y42A5A.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Steward C.A.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RX SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology.";
RL Science 282:2012-2018(1998).
CC -|- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AL032618; CAB63367.1; -.
DR HSSP; P24941; LHCL.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_kin_actsite.
DR Pfam; PF00069; pkinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 353 AA; 40665 MW; 7E32260D9849CAC9 CRC64;

Query Match 61.8%; Score 1027.5; DB 5; Length 353;
Best Local Similarity 58.8%; Pred. No. 8.9e-83;
Matches 187; Conservative 59; Mismatches 63; Indels 9; Gaps 2;

Query Match		46.3%;	Score 769.5;	DB 5;	Length 1106;				
Best Local Similarity		48.3%;	Pred. No. 2.7e-59;						
Matches 145;		Conservative 62;	Mismatches 88;	Indels 5;	Gaps 2;				
QY	1	MEKYEKLAKTGE	SGYGVVFKCRNKTSGQVVAVKKFVESEDDPVVKKIALREIRMLKQLKH	60					
Db	1	MEAYETLGILGEG	TYGVVVKARSVRTGKLVAIKRFQTEQDEHVRKTSREVRMLQLQH	60					
QY	61	PNLVNLIIEVFR	RRKRMHLVFEYCDHTLLNELERNPNGVADGVIKSVLWQTLQALNFC	120					
Db	61	PNVIRLEDVFR	REGKLYLVFEFIDHTILQLLESTTRGFRHRELRRYTYQLLRGIEFCH	120					
QY	121	NCIHRDIKPEN	ILITKOGIIKICDFGFAQILIPGDATDYVATRWYRAPELLVGD	180					
Db	121	NIHRDVKPEN	VLIDESGLLKLCDFGFAQTSAGKGYTDYVATRWYRAPELLVGD	180					
QY	181	SVDIWAIGCV	FAELLTGQPLWPKSDVDQLYLIIRTLGKLIPRHQSIFKSN	240					
Db	181	PVDWALGCM	FAELSDGQPLFGESDLQCLIMQTCGPVPQRLVFI	240					
QY	241	EPEDMETLEE	KFSDVHPVALNFMKGCLKNPDDRLTCSQLLESSYF--DSFQ--	295					
Db	241	HTDILYTLK	DRYHRESNDWIEFLSSCLHTDPAQRLCTELMELPYFTRDGRDRY	300					
RESULT 11									
Q9JM02									
ID	Q9JM02	PRELIMINARY;		PRT;	457 AA.				
AC	Q9JM02;								
DT	01-OCT-2000	(TrEMBLrel. 15, Created)							
DT	01-OCT-2000	(TrEMBLrel. 15, Last sequence update)							
DT	01-JUN-2001	(TrEMBLrel. 17, Last annotation update)							
DE	SERINE/THREONINE KINASE NKIATRE BETA.								
OS	Rattus norvegicus (Rat).								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.								
OX	NCBI_TaxID=10116;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RA	Haq R., Randall S., Midmer M., Iafrate B., Zanke B.W.;								
RT	"Cloning and functional characterization of NKIATRE, a novel kinase								
RT	related to both MAPKs and cyclin-dependent kinases.";								
RL	Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.								
CC	1- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.								
DR	EMBL; AF112183;	AAF34870.1; -							
DR	InterPro; IPR000719; Euk_pkinase.								
DR	InterPro; IPR002290; Ser_thr_kin_actsite.								
DR	Pfam; PF00069; pkinase; 1.								
DR	SMART; SM00220; S_TKC; 1.								
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.								
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.								
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.								
KW	ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.								
SQ	SEQUENCE 457 AA; 51894 MW; 99D0782794099CDA CRC64;								
Query Match 45.9%; Score 763.5; DB 11; Length 457;									
Best Local Similarity		46.9%;	Pred. No. 2.7e-59;						
Matches 145;		Conservative 65;	Mismatches 92;	Indels 7;	Gaps 4;				
QY	1	MEKYEKLAKTGE	SGYGVVFKCRNKTSGQVVAVKKFVESEDDPVVKKIALREIRMLKQLKH	60					
Db	1	MEMYETLGKVG	EGSYGVVMCKKHKTGRIVAIIKIFYE-KPEKSVNKIATREIKELKQFRH	59					
QY	61	PNLVNLIIEVFR	RRKRMHLVFEYCDHTLLNELERNPNGVADGVIKSVLWQTLQALNFC	120					
Db	60	ENLVNLIIEVFR	QKKTHLVFEFIDHTVLDELQHYCHGLESKRRLKRYLFQILRAIEY	119					
Query Match 45.9%; Score 763.5; DB 11; Length 457;									
Best Local Similarity		46.9%;	Pred. No. 2.7e-59;						
Matches 145;		Conservative 65;	Mismatches 92;	Indels 7;	Gaps 4;				
QY	1	MEKYEKLAKTGE	SGYGVVFKCRNKTSGQVVAVKKFVESEDDPVVKKIALREIRMLKQLKH	60					
Db	1	MEMYETLGKVG	EGSYGVVMCKKHKTGRIVAIIKIFYE-KPEKSVNKIATREIKELKQFRH	59					
QY	61	PNLVNLIIEVFR	RRKRMHLVFEYCDHTLLNELERNPNGVADGVIKSVLWQTLQALNFC	120					
Db	60	ENLVNLIIEVFR	QKKTHLVFEFIDHTVLDELQHYCHGLESKRRLKRYLFQILRAIEY	119					
QY	121	NCIHRDIKPEN	ILITKOGIIKICDFGFAQIL-IPGDATDYVATRWYRAPELLVGD	179					
Db	120	NIHRDIKPEN	ILVSQGITKLCDFGFARTLAAPGDVTDYVATRWYRAPELLVKD	179					

QY	180	SSVDIWAIGCVFAELLTGQPLWPKSDVDQLYLIIRTLGKLIPRHQSIFKSN	239	
Db	180	KPVDIWAIGCMIIEMATGNPYLPSSDDLHLHKIVLVGNLTPLHLHNIFSKSPIFAGVVL	239	
QY	240	PEPEDMETLEEKFSVDHPVALNFMKGCLKNPDDRLTCSQLLESSYF--DSFQE---AQI	294	
Db	240	PQVQHPKNARKKYPKLNGLLADIHVHACIQIDPAERISSTDLLHHDYFTRDGFIEKFIPEL	299	
QY	295	KRKARNEGR 303		
Db	300	RAKLLQEA 308		
RESULT 12				
Q9JM01				
ID	Q9JM01	PRELIMINARY;	PRT; 505 AA.	
AC	Q9JM01;			
DT	01-OCT-2000	(TrEMBLrel. 15, Created)		
DT	01-OCT-2000	(TrEMBLrel. 15, Last sequence update)		
DT	01-JUN-2001	(TrEMBLrel. 17, Last annotation update)		
DE	SERINE/THREONINE KINASE NKIATRE ALPHA.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Haq R., Randall S., Midmer M., Iafrate B., Zanke B.W.;			
RT	"Cloning and functional characterization of NKIATRE, a novel kinase			
RT	related to both MAPKs and cyclin-dependent kinases.";			
RL	Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.			
CC	-!- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.			
DR	EMBL; AF112184;	AAF34871.1; -		
DR	InterPro; IPR000719; Euk_pkinase.			
DR	InterPro; IPR002290; Ser_thr_kin_actsite.			
DR	Pfam; PF00069; pkinase; 1.			
DR	SMART; SM00220; S_TKc; 1.			
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.			
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.			
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.			
KW	ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.			
SQ	SEQUENCE 505 AA; 57242 MW; 89747921FF2CE683 CRC64;			
Query Match 45.9%; Score 763.5; DB 11; Length 505;				
Best Local Similarity 46.9%; Pred. No. 3.1e-59;				
Matches 145; Conservative 65; Mismatches 92; Indels 7; Gaps 4;				
QY	1	MEKYEKLAKTGE	SGYGVVFKCRNKTSGQVVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60	
Db	1	MEMYETLGKVG	EGSYGVVMCKKHKTGRIVAIIKIFYE-KPEKSVNKIATREIKELKQFRH 59	
QY	61	PNLVNLIIEVFR	KRMHLVFEYCDHTLLNLELRNPNGVADGVIKSVLWQTLQALNFC	120
Db	60	ENLVNLIIEVFR	QKKTHLVFEFIDHTVLDELQHYCHGLESKRRLKRYLFQILRAIEY	119
QY	121	NCIHRDIKPEN	ILITKOGIIKICDFGFAQIL-IPGDATDYVATRWYRAPELLVGD	179
Db	120	NIHRDIKPEN	ILVSQGITKLCDFGFARTLAAPGDVTDYVATRWYRAPELLVKD	179
QY	180	SSVDIWAIGCVFAELLTGQPLWPKSDVDQLYLIIRTLGKLIPRHQSIFKSN	239	
Db	180	KPVDIWAIGCMIIEMATGNPYLPSSDDLHLHKIVLVGNLTPLHLHNIFSKSPIFAGVVL	239	
QY	240	PEPEDMETLEEKFSVDHPVALNFMKGCLKNPDDRLTCSQLLESSYF--DSFQE---AQI	294	
Db	240	PQVQHPKNARKKYPKLNGLLADIHVHACIQIDPAERISSTDLLHHDYFTRDGFIEKFIPEL	299	
QY	295	KRKARNEGR 303		
Db	300	RAKLLQEA 308		

Fri May 3 11:22:28 2002

GN	STK9.	OS	Homo sapiens (Human).	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC		OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	OX	NCBI_TaxID=9606;
RN	[1]	RP	SEQUENCE FROM N.A.	RA	Bird C.;
RL	Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.	CC	-!- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.	DR	EMBL; Z92542; CAB38870.1; -.
DR	HSSP; P28523; 1A60.	DR	InterPro; IPR000719; Euk_pkinase.	DR	InterPro; IPR000508; Peptidase_S26.
DR	InterPro; IPR002290; Ser_thr_kin_actsite.	DR	Pfam; PF00069; pkinase; 1.	DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.	DR	PROSITE; PS00501; SPASE_1_1; UNKNOWN_1.	KW	ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
FT	NON_TER 1	SQ	SEQUENCE 997 AA; 111939 MW; B922855AED72608D CRC64;		
Query Match 38.6%; Score 641.5; DB 4; Length 997;					
Best Local Similarity 43.4%; Pred. No. 4.8e-48;					
Matches 124; Conservative 65; Mismatches 90; Indels 7; Gaps 4;					
Qy	28	QVAVKKEVESEDDPVVKKIALREIRMLKQLKHPNLVNLIEVFRKRKMHLVFEYCDHTL	87		
Db	4	EIVAIKFKDSEENEVEKETTRELKMLRTLKQENIVELKEAFRRRGKLYLVFEYVEKNM	63		
Qy	88	LNELEARNPENGVDGVKSVLWQTLOALNFCIHNCIHRDIKIPENILITKQGIKIDFGF	147		
Db	64	LELLEMPNGVPPEKVKSYIYQLIKAIHWCHKNDIVHRDIKIPENLLISHNDVLKCDFGF	123		
Qy	148	AQILIPGD--AYDYVATRWYRAPELLVGDYQYGVSSVDIWAIGCVFAELLTGQPLWPGKS	205		
Db	124	ARNLSEGNANYTEYVATRWYRSPPELLLG-APYKSVDMWSVGCILGELSDGQPLFPGES	182		
Qy	206	DVDQYLYLIIRTLGKLIIPRHQSIFKSNGFHGISIPEPEDMETLEEKFSDV-HPVALNFMK	264		
Db	183	EIDQLFTIQKVLGPLPAEQMKLFYNPRFHLRPPAVNHPQSLRRLYLGLINSVLLDLMK	242		
Qy	265	GCLKMNPDDRLTCSQLLESSYFDSFQEAQIKRKARNEGRNRRQQV	310		
Db	243	NLLKLDPADRYLTEQCLN---HPTFQTQRLDRSPSRSAKRKPVHV	285		
RESULT 15					
Q9ZRI1		PRELIMINARY;	PRT;	294	AA.
ID	Q9ZRI1				
AC	Q9ZRI1;				
DT	01-MAY-1999 (TrEMBLrel. 10, Created)				
DT	01-MAY-1999 (TrEMBLrel. 10, Last sequence update)				
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)				
DE	P34CDC2.				
GN	CDC2TAA.				
OS	Triticum aestivum (Wheat).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;				
OC	Triticeae; Triticum.				
OX	NCBI_TaxID=4565;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Dong C., John P.C.L.;				
RL	Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.				
CC	-!- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.				
DR	EMBL; U23409; AAD10483.1; -.				
DR	HSSP; P24941; 1JSU.				
DR	Mendel; 37149; Triae;2321;37149.				
DR	InterPro; IPR000719; Euk_pkinase.				
DR	InterPro; IPR002290; Ser_thr_kin_actsite.				

RESULT 13					
Q9W6R6		PRELIMINARY;	PRT;	1104	AA.
ID	Q9W6R6				
AC	Q9W6R6;				
DT	01-NOV-1999 (TrEMBLrel. 12, Created)				
DT	01-NOV-1999 (TrEMBLrel. 12, Last sequence update)				
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)				
DE	SERINE-THREONINE KINASE 9.				
GN	STK9.				
OS	Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;				
OC	Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;				
OC	Tetraodontidae; Takifugu.				
OX	NCBI_TaxID=31033;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=99263230; PubMed=10330123;				
RA	Brunner B., Todt T., Lenzner S., Stout K., Schulz U., Ropers H.-H.,				
RA	Kalscheuer V.M.;				
RT	"Genomic structure and comparative analysis of nine Fugu genes:				
RT	conservation of synteny with human chromosome Xp22.2-p22.1.";				
RL	Genome Res. 9:437-448(1999).				
CC	-!- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.				
DR	EMBL; AF146687; AAD28798.1; -.				
DR	HSSP; Q16539; 1WFC.				
DR	InterPro; IPR000719; Euk_pkinase.				
DR	InterPro; IPR002290; Ser_thr_kin_actsite.				
DR	Pfam; PF00069; pkinase; 1.				
DR	SMART; SM00220; S_TKc; 1.				
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.				
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.				
KW	ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.				
SQ	SEQUENCE 1104 AA; 123751 MW; 9085A078B551D876 CRC64;				
Query Match 38.7%; Score 644; DB 13; Length 1104;					
Best Local Similarity 42.0%; Pred. No. 3.3e-48;					
Matches 128; Conservative 66; Mismatches 97; Indels 14; Gaps 6;					
Qy	24	KTSGQVAVKKEVESEDDPVVKKIALREIRMLKQLKHPNLVNLIEVFRKRKMHLVFEYC	83		
Db	19	KETNELVAIKKFKDSEENEVEKETTRELKMLRTLKQDNIVELKEAFRRRGKLYLVFEYV	78		
Qy	84	DHTLLNELEARNPENGVDGVKSVLWQTLOALNFCIHNCIHRDIKIPENILITKQGIKIC	143		
Db	79	ERNMELLEHEPTGAPPDKVRYSYIYQLIKAINWCHKNEIVHRDIKIPENLLISSDDILKLC	138		
Qy	144	DGFAQILIPG-DA-YDYVATRWYRAPELLVGDYQYGVSSVDIWAIGCVFAELLTGQPLW	201		
Db	139	DGFAFARNLSEGTDANYTEYVATRWYRSPPELLLG-APYKAVDMWSVGCILGELSDGQPLF	197		
Qy	202	PGKSDVDQYLYLIIRTLGKLIIPRHQSIFKSNGFHGISIPEPEDMETLEEKFSDV-HPVAL	260		
Db	198	PGESEIDQLFTIQKVLGPLPAEQMKLFYNPRFHLRPPAVNHPQSLRRLYLGLINSGLML	257		
Qy	261	NFMKGCLKMNPDDRLTCSQLLESSYFDSFQEAQIKRKA---RNEGRNRR-----QQV	310		
Db	258	DLMKNLLLNPTERYLTEQSLNHPAFQPLQROVERERAPPASPNDPRSSKRKTHHHGENT	317		
Qy	311	LPLKS 315			
Db	318	VPTRS 322			
RESULT 14					
Q9UJL6		PRELIMINARY;	PRT;	997	AA.
ID	Q9UJL6				
AC	Q9UJL6;				
DT	01-MAY-2000 (TrEMBLrel. 13, Created)				
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)				
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)				
DE	DJ245G19.2 (SERINE/THREONINE KINASE 9) (FRAGMENT).				

pfam: PF00069; pkinase; 1.
 SMART; SM00220; S_TKc; 1.
 PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 294 AA; 33996 MW; 77ABC487FA862A72 CRC64;

Query Match 34.9%; Score 579.5; DB 10; Length 294;
 Best Local Similarity 38.2%; Pred. No. 2.8e-43;
 Matches 117; Conservative 71; Mismatches 93; Indels 25; Gaps

QY	1	MEKYEK	LAKT	GGSGYGVV	EKCRNKTSGQVAVK	KFVESEDDPVVK	KKIALREIRMLKQ	LKH 60
Dd	1	MEQYK	VEKIGEGTYGVVYK	ARDRTTNETIAL	KIRLEQ	EGDEGVPSTAIR	ISLLKEMQH 60	
QY	61	PNLVN	LIEVFR	RRKRMHLV	FEYCDHTLL	NELEARNP	-GVADGVIKSVLWQTLQALNFCHI 119	
Dd	61	GNIVK	LHDVVHSEKRIW	LVFEYLDL	LKKFMDSCPEFAKSPALIKSYLQILRGVAYCHS 120			
QY	120	HNCIH	ROI	KPENILITKQ-G	LIKICDFGAQIL-IPGDAYTDYVATRWYRAPELLVGDTQ 177			
Dd	121	HRVLH	RDLK	QPQLLIDRR	TNALKLADFLARAF	IGPVRTFTEHVVTLWYRAPELLIGARQ 180		
QY	178	YGSSV	DIWAIGCVFAELL	TCQPLWPGKSDVDQ	LVLIIRTLGKLIPRHQSIFKSNCFEHGI 237			
Dd	181	YSTP	VDWMSVGCIFAEMV	NQKPLPFGDSEIDELFKIFRVLG--TPNEQT-----WPGV 231				
QY	238	-SIP-----	EPEDMETLEEK	FSDVHPVALNFMKGC	LKMNPDDRLTCSOLLESSYFD 287			
Dd	232	SSLPDYKSAFPRWQA	EDLATV---VPNLEPVG	LDLLSKMLRFEPNKRITARQALEHEYFK 288				
QY	288	SFQEAQ	293					
Dd	289	DMEVMQ	294					

Search completed: May 3, 2002, 11:08:13
Job time: 352 sec

us-09-671-050-12.ra1

Fri May 3 11:22:27 2002

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2002, 11:02:44 ; Search time 22.35 Seconds
(without alignments)
317.160 Million cell updates/sec

Title: US-09-671-050-12
Perfect score: 1662
Sequence: 1 MEKYEKLAKTGEYSYGVFK.....RKARNEGRRRRQQVLPLKS 315

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	570	34.3	298	2 US-08-874-347-25	Sequence 25, Appl
2	570	34.3	298	2 US-08-969-106-2	Sequence 2, Appli
3	570	34.3	298	3 US-09-093-522-25	Sequence 25, Appl
4	568.5	34.2	294	2 US-08-874-347-26	Sequence 26, Appl
5	568.5	34.2	294	3 US-09-093-522-26	Sequence 26, Appl
6	565	34.0	297	1 US-08-176-620A-16	Sequence 16, Appl
7	565	34.0	297	2 US-08-461-985-16	Sequence 16, Appl
8	557	33.5	297	2 US-08-874-347-23	Sequence 23, Appl
9	557	33.5	297	3 US-09-093-522-23	Sequence 23, Appl
10	556	33.5	297	2 US-08-874-347-24	Sequence 24, Appl
11	556	33.5	297	3 US-09-093-522-24	Sequence 24, Appl
12	555.5	33.4	300	2 US-08-874-347-10	Sequence 10, Appl
13	555.5	33.4	300	3 US-09-093-522-10	Sequence 10, Appl
14	555	33.4	270	2 US-07-857-224B-31	Sequence 31, Appl
15	551	33.2	274	1 US-08-318-947A-20	Sequence 20, Appl
16	551	33.2	274	2 US-08-795-303-20	Sequence 20, Appl
17	544	32.7	292	1 US-08-154-915-2	Sequence 2, Appli
18	544	32.7	292	2 US-08-464-517-38	Sequence 38, Appl
19	544	32.7	292	2 US-08-246-361A-38	Sequence 38, Appl
20	544	32.7	292	3 US-08-463-772-38	Sequence 38, Appl
21	544	32.7	292	5 PCT-US93-09945-2	Sequence 2, Appli
22	537	32.3	297	4 US-08-932-787B-21	Sequence 21, Appl
23	537	32.3	297	4 US-08-932-012C-21	Sequence 21, Appl
24	537	32.3	297	4 US-08-888-818C-21	Sequence 21, Appl
25	518.5	31.2	297	2 US-08-874-347-22	Sequence 22, Appl
26	518.5	31.2	297	3 US-09-093-522-22	Sequence 22, Appl
27	511	30.7	274	2 US-07-857-224B-30	Sequence 30, Appl

28	505.5	30.4	317	1	US-08-463-090B-9	Sequence 9, Appli
29	505.5	30.4	317	2	US-08-874-347-18	Sequence 18, Appl
30	505.5	30.4	317	3	US-09-093-522-18	Sequence 18, Appl
31	505	30.4	323	2	US-08-874-347-21	Sequence 21, Appl
32	505	30.4	323	3	US-09-093-522-21	Sequence 21, Appl
33	500.5	30.1	282	1	US-08-318-947A-19	Sequence 19, Appl
34	500.5	30.1	282	2	US-08-795-303-19	Sequence 19, Appl
35	496	29.8	324	2	US-08-874-347-20	Sequence 20, Appl
36	496	29.8	324	3	US-09-093-522-20	Sequence 20, Appl
37	485	29.2	274	2	US-07-857-224B-29	Sequence 29, Appl
38	484	29.1	316	1	US-08-403-634-4	Sequence 4, Appli
39	484	29.1	316	4	US-08-913-441B-4	Sequence 4, Appli
40	482.5	29.0	326	2	US-08-306-511A-10	Sequence 10, Appl
41	482.5	29.0	326	2	US-08-893-274-10	Sequence 10, Appl
42	481.5	29.0	298	2	US-08-061-636-3	Sequence 3, Appli
43	481.5	29.0	298	2	US-08-874-347-19	Sequence 19, Appl
44	481.5	29.0	298	3	US-09-093-522-19	Sequence 19, Appl
45	481.5	29.0	298	5	PCT-US94-05268-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-874-347-25
; Sequence 25, Application US/08874347
; Patent No. 5863741
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; TITLE OF INVENTION: CARINII
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/874,347
; FILING DATE: 13-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/055001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070
; TELEFAX: 612-288-9696
; TELEX:
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-874-347-25

Query Match 34.3%; Score 570; DB 2; Length 298;
Best Local Similarity 40.0%; Pred. No. 5.5e-52;
Matches 118; Conservative 60; Mismatches 99; Indels 18; Gaps 6;

Db 61 GNVRLHVDVIHSEKRIYLVFEYLDLDLKKFMDSCPEFAKNPT-----LIKSYLYQILRGV 115
QY 115 NFCHIHNCIHRDIKPENILITKQ-GIIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELL 172
Db 116 AYCHSHRVLHRDLKPQNLIDRRTNALKLADFGLARAFGIPVRTFHEVVTLWYRAPEIL 175
QY 173 VGDTQYGGSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIPRHQSIFKSNG 232
Db 176 LGSROYSTPVDWMSVGCIFAEMVNQKPLFPGDSEIDELFKIFRVLG--TPNEQS----- 227
QY 233 FFHGI-SIP-----EPEDMETLEEKFSVDHPVALNFMKGCLKMNPDDRLTCSQILLE 282
Db 228 -WPGVSSLPDYKSAFPKWQAQDLATI---VPTLDPAGLDLLSKMLRYEPNKRITARQALE 283
QY 283 SSYFDSFQEAQ 293
Db 284 HEYFKDLEMVQ 294

RESULT 5
US-09-093-522-26
; Sequence 26, Application US/09093522
; Patent No. 6015700
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; TITLE OF INVENTION: CARINII
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/093,522
; FILING DATE: 08-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/874,347
; FILING DATE: 13-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/055002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070
; TELEFAX: 612-288-9696
; TELEX:
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 294 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-093-522-26

Query Match 34.2%; Score 568.5; DB 3; Length 294;
Best Local Similarity 37.6%; Pred. No. 7.8e-52;
Matches 117; Conservative 67; Mismatches 92; Indels 35; Gaps 9;

QY 61 PNLVNLIIEVFRKRKMHLVFEYCDHTLLNELEARNP-NGVADGVYKSVLWQTLQALNFCHI 119
Db 61 PNIIVKLLDVHTENKLYLVFEFLHQDLKKFMDASALTGIPLPLIKSYLFQLLQGLAFCHS 120
QY 120 HNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDTOY 178
Db 121 HRVLHRDLKPQNLINTEGAIKLADFGLARAFGVPVRTYHEVVTWYRAPEILLGSKYY 180
QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIPRHQSIFKSNGFFHGI- 237
Db 181 STAVDIWSLGCIFAEMVTRRALFPGDSEIDQLFRIFRTLGT-----PDEVVWPGVT 231
QY 238 SIPE--PEDMETLEEKFSVDHPV-----ALNFMKGCLKMNPDDRLTCSQILLESSYF 286
Db 232 SMPDYKPSFPKWARQDFSKVVPPLDEDDGRSLLSQMLHYDPNKRISAKAALAHPPF 286

RESULT 4
US-08-874-347-26
; Sequence 26, Application US/08874347
; Patent No. 5863741
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; TITLE OF INVENTION: CARINII
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/874,347
; FILING DATE: 13-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/055001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070
; TELEFAX: 612-288-9696
; TELEX:
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 294 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-874-347-26

Query Match 34.2%; Score 568.5; DB 2; Length 294;
Best Local Similarity 37.6%; Pred. No. 7.8e-52;
Matches 117; Conservative 67; Mismatches 92; Indels 35; Gaps 9;

QY 1 MEKYEKLAKTGEGSYGVVFKCRNKTSQGVVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEQYEKEKIGEGTYGVYRARDKVTNETIALKIRLEQDEGVPSTAIRAISLLKEMHH 60
QY 61 PNLVNLIIEVFRKRKMHLVFEYCDHTLLN-----ELERNPNGVADGVYKSVLWQTLQAL 114

QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSQVVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEQYEKEEKIGEGTYGVVYRARDKVTNETIALKKIRLQEQEDEGVPTAIREISLLKEMHH 60
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLN-----ELERNPNGVADGVIKSVLWQTLQAL 114
Db 61 GNIVRLHDVHSEKRIYLVFEYLDLDLKKFMDSCPEFAKNPT-----LIKSVLYQILRGV 115
QY 115 NFCHIHNCIHRDIPENILITKQ-GIITKIDFGFAQIL-IPGDAYTDYVATRWYRAPELL 172
Db 116 AYCHSHRVLHRDLKPQNLLIDRRTNALADFGLARAFGIPVRTFTHEVVTLWYRAPEIL 175
QY 173 VGDQYSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIPRHQSIFKSNG 232
Db 176 LGSROYSTPVDMSVGCIFAEMVNQKPLFPGDSEIDELFKIFRVLG--TPNEQS----- 227
QY 233 FFHGI-SIP-----EPEDMETLEEKFSVDHPVALNFMKGCLKMNPDORLTCSQLLE 282
Db 228 -WPGVSSLPDYKSAPFKWQAQDLATI---VPTLDPAGLDLISKMLRYEPNKRITARQALE 283
QY 283 SSYFDSFQEAQ 293
Db 284 HEYFKDLEVMVQ 294

RESULT 6
US-08-176-620A-16
; Sequence 16, Application US/08176620A
; Patent No. 5595904
; GENERAL INFORMATION:
; APPLICANT: Boulton, Teri G.
; APPLICANT: Cobb, Melanie H.
; APPLICANT: Yancopoulos, George D.
; APPLICANT: Nye, Steven
; APPLICANT: Panayotatos, Nikos
; TITLE OF INVENTION: A Family of Map2 Protein Kinases
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/176,620A
; FILING DATE: 03-JAN-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 6526-123
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 297 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-176-620A-16

Query Match 34.0%; Score 565; DB 1; Length 297;
Best Local Similarity 39.7%; Pred. No. 1.8e-51;

Matches 124; Conservative 61; Mismatches 95; Indels 32; Gaps 8;
QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSQVVAVKKF-VESEDDPVVKKIALREIRMLKQLK 59
Db 1 MEDYTKIEKIGEGTYGVVYKGRHKTTGQVVAMKKIRLESEEG-VPSTAIRESILLKELR 59
QY 60 HPNLVNLIIEVFRKRKMHVFEYCDHTLILNELEARNPG--VADGVIKSVLWQTLQALNFC 117
Db 60 HPNIVSLQDVLMDQDSRLYLIFEFLSMDLKKYLDISIPPGQYMDSSLVKSYLEYQILQGIIVFC 119
QY 118 HIHNCIHRDIKPENILITKOGIIKICIDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGD 176
Db 120 HSRRLVLRDLKPQNLLIDDKGTIKLADFGLARAFGIPRVYTHVVVTLWYRSPEVLLGSA 179
QY 177 QYSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIPRHQSIFKSNGFFHG 236
Db 180 RYSTPVDIWSIGTIFAELATKKPLFHGDSEIDQLFRIFRALG--TPNNE----- 226
QY 237 ISPEPEDMETLEEKFSVDHPVA-----LNFMKGCLKMNPDORLTCSQLLESS 284
Db 227 -VMPEVESLQDYKNTFPKWKPGSLASHVKNLDENGLDLSKMLIYDPAKRISGKMALNHP 285
QY 285 YFDSFQEAQIKR 296
Db 286 YFNDL-DNQIKK 296

RESULT 7
US-08-461-985-16
; Sequence 16, Application US/08461985
; Patent No. 5872006
; GENERAL INFORMATION:
; APPLICANT: Boulton, Teri G.
; APPLICANT: Cobb, Melanie H.
; APPLICANT: Yancopoulos, George D.
; APPLICANT: Nye, Steven
; APPLICANT: Panayotatos, Nikos
; TITLE OF INVENTION: A Family of Map2 Protein Kinases
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,985
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/176,620
; FILING DATE: 03-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 6526-123
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 297 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein

US-08-461-985-16

Query Match 34.0%; Score 565; DB 2; Length 297;
Best Local Similarity 39.7%; Pred. No. 1.8e-51;
Matches 124; Conservative 61; Mismatches 95; Indels 32; Gaps 8;

QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVAVKKF-VESEDDPVVKKIALREIRMLKQLK 59
Db 1 MEDYTKIEKIGEGTYGVVYKGRHRTTGQIVAMKKIRLESEEG-VPSTAIRESILLKELR 59

QY 60 HPNLVNLIEVFRRRKRMHLVFEYCDHTLLNELEPNPNG--VADGVIKSVLWOTLQALNFC 117
Db 60 HPNIVSLQDVLMDQSRLLYLIFEFLSMDLKKYLDLSDIPPGQFMDSSLVKSILYQILOGIVFC 119

QY 118 HHNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELIVGDT 176
Db 120 HSRRVLHRLDKPQNLLIDDKGTIKLADFGLARAFGIPRVYTHYVVTWYRSPEVLLGSA 179

QY 177 QYGSSVDIWAIGCVFAELLTGQPLWPGKSDVDOLYLIIRTGLKLIIPRHQSIFKSNFFHG 236
Db 180 RYSTPVDIWSIGTIFAELATKKPLFHGDSEIDQLFRIFRALTG--TPNNE----- 226

QY 237 ISIPEPEMETLEEKFSVHPVA-----LNFMKGCLKMNPDRLTCSQLLESS 284
Db 227 -VWPEVESLQDYKNTFPKWKPGSLASHVKNLDEGLDLSKMLIYDPAKRISGKMALNHP 285

QY 285 YFDSFQEAQIKR 296
Db 286 YFNDL-DNQIKK 296

US-08-874-347-23

Query Match 33.5%; Score 557; DB 2; Length 297;
Best Local Similarity 39.1%; Pred. No. 1.3e-50;
Matches 122; Conservative 63; Mismatches 95; Indels 32; Gaps 8;

QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVAVKKF-VESEDDPVVKKIALREIRMLKQLK 59
Db 1 MEDYTKIEKIGEGTYGVVYKGRHRTTGQIVAMKKIRLESEEG-VPSTAIRESILLKELR 59

QY 60 HPNLVNLIEVFRRRKRMHLVFEYCDHTLLNELEPNPNG--VADGVIKSVLWOTLQALNFC 117
Db 60 HPNIVSLQDVLMDQSRLLYLIFEFLSMDLKKYLDLSDIPPGQFMDSSLVKSILYQILOGIVFC 119

QY 118 HHNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELIVGDT 176
Db 120 HSRRVLHRLDKPQNLLIDDKGTIKLADFGLARAFGIPRVYTHYVVTWYRSPEVLLGSA 179

QY 177 QYGSSVDIWAIGCVFAELLTGQPLWPGKSDVDOLYLIIRTGLKLIIPRHQSIFKSNFFHG 236
Db 180 RYSTPVDIWSIGTIFAELATKKPLFHGDSEIDQLFRIFRALTG--TPNNE----- 226

QY 237 ISIPEPEMETLEEKFSVHPVA-----LNFMKGCLKMNPDRLTCSQLLESS 284
Db 227 -VWPEVESLQDYKNTFPKWKPGSLASHVKNLDEGLDLSKMLVYDPAKRISGKMALKHP 285

QY 285 YFDSFQEAQIKR 296
Db 286 YFDDL-DNQIKK 296

RESULT 9

US-09-093-522-23

; Sequence 23, Application US/09093522
; Patent No. 6015700

; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; TITLE OF INVENTION: CARINII
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/093,522
; FILING DATE: 08-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/874,347
; FILING DATE: 13-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/055002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070
; TELEFAX: 612-288-9696
; TELEX:

US-08-874-347-23

Query Match 34.0%; Score 565; DB 2; Length 297;
Best Local Similarity 39.7%; Pred. No. 1.8e-51;
Matches 124; Conservative 61; Mismatches 95; Indels 32; Gaps 8;

QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVAVKKF-VESEDDPVVKKIALREIRMLKQLK 59
Db 1 MEDYTKIEKIGEGTYGVVYKGRHRTTGQIVAMKKIRLESEEG-VPSTAIRESILLKELR 59

QY 60 HPNLVNLIEVFRRRKRMHLVFEYCDHTLLNELEPNPNG--VADGVIKSVLWOTLQALNFC 117
Db 60 HPNIVSLQDVLMDQSRLLYLIFEFLSMDLKKYLDLSDIPPGQFMDSSLVKSILYQILOGIVFC 119

QY 118 HHNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELIVGDT 176
Db 120 HSRRVLHRLDKPQNLLIDDKGTIKLADFGLARAFGIPRVYTHYVVTWYRSPEVLLGSA 179

QY 177 QYGSSVDIWAIGCVFAELLTGQPLWPGKSDVDOLYLIIRTGLKLIIPRHQSIFKSNFFHG 236
Db 180 RYSTPVDIWSIGTIFAELATKKPLFHGDSEIDQLFRIFRALTG--TPNNE----- 226

QY 237 ISIPEPEMETLEEKFSVHPVA-----LNFMKGCLKMNPDRLTCSQLLESS 284
Db 227 -VWPEVESLQDYKNTFPKWKPGSLASHVKNLDEGLDLSKMLIYDPAKRISGKMALNHP 285

QY 285 YFDSFQEAQIKR 296
Db 286 YFNDL-DNQIKK 296

US-08-874-347-23

Query Match 34.0%; Score 565; DB 2; Length 297;
Best Local Similarity 39.7%; Pred. No. 1.8e-51;
Matches 124; Conservative 61; Mismatches 95; Indels 32; Gaps 8;

QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVAVKKF-VESEDDPVVKKIALREIRMLKQLK 59
Db 1 MEDYTKIEKIGEGTYGVVYKGRHRTTGQIVAMKKIRLESEEG-VPSTAIRESILLKELR 59

QY 60 HPNLVNLIEVFRRRKRMHLVFEYCDHTLLNELEPNPNG--VADGVIKSVLWOTLQALNFC 117
Db 60 HPNIVSLQDVLMDQSRLLYLIFEFLSMDLKKYLDLSDIPPGQFMDSSLVKSILYQILOGIVFC 119

QY 118 HHNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELIVGDT 176
Db 120 HSRRVLHRLDKPQNLLIDDKGTIKLADFGLARAFGIPRVYTHYVVTWYRSPEVLLGSA 179

QY 177 QYGSSVDIWAIGCVFAELLTGQPLWPGKSDVDOLYLIIRTGLKLIIPRHQSIFKSNFFHG 236
Db 180 RYSTPVDIWSIGTIFAELATKKPLFHGDSEIDQLFRIFRALTG--TPNNE----- 226

QY 237 ISIPEPEMETLEEKFSVHPVA-----LNFMKGCLKMNPDRLTCSQLLESS 284
Db 227 -VWPEVESLQDYKNTFPKWKPGSLASHVKNLDEGLDLSKMLIYDPAKRISGKMALNHP 285

QY 285 YFDSFQEAQIKR 296
Db 286 YFNDL-DNQIKK 296

RESULT 8

US-08-874-347-23

; Sequence 23, Application US/08874347
; Patent No. 5863741

; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; TITLE OF INVENTION: CARINII
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/874,347
; FILING DATE: 13-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/055001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070
; TELEFAX: 612-288-9696
; TELEX:
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 297 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

us-09-671-050-12.ra1

Fri May 3 11:22:27 2002

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ellinger, Mark S.
REGISTRATION NUMBER: 34,812
REFERENCE/DOCKET NUMBER: 07039/055001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-335-5070
TELEFAX: 612-288-9696
TELEX:
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 300 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-874-347-10

Query Match 33.4%; Score 555.5; DB 2; Length 300;
Best Local Similarity 37.3%; Pred. No. 1.9e-50;
Matches 113; Conservative 72; Mismatches 91; Indels 27; Gaps 8;
QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEDYTKIEKIGEGTYGVVYKGRHKTTGQVAMKKIRLESEEEG-VPSTAIRISLLKELR 59
QY 61 HNLVNLIEVFRKRKMHVFEYCDHTLLNELELRNPNGVADG--VIKSVLWQTLQALNFC 118
Db 61 DNVVRLNIIHQESRLYLVEFDLDLKKYMNISIPKDMMLGAEMIKKFMSQLVSGVKYCH 120
QY 119 IHNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGD 177
Db 121 SHRILHRDLKPQNLLIDREGNLKLADFLARAFGVPLRGYTHEVVTLWYRAPEVLLGGR 180
QY 178 YGSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIPIRHQSIFKSN 237
Db 181 YATLDIWSIGCIFAEAMATKPKPLFGDSEIDEIFRIFRILG--TPDENS-----WPGI 231
QY 238 -SIPE-----PEDM-ETLEKFSVDVHPVALNFMKGLKMNPPDDRLTCSQLLESSYF 286
Db 232 TSYPDFKATFPKWSPNKLGELITELSDSD----GIDLLQCLRYYP AERISAKKALDHPYF 287
QY 287 DSF 289
Db 288 DDF 290

RESULT 13
US-09-093-522-10
; Sequence 10, Application US/09093522
; Patent No. 6015700
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; TITLE OF INVENTION: CARINII
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for windows Version 2.0

NAME: Ellinger, Mark S.
REGISTRATION NUMBER: 34,812
REFERENCE/DOCKET NUMBER: 07039/055002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-335-5070
TELEFAX: 612-288-9696
TELEX:
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 297 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-093-522-24

Query Match 33.5%; Score 556; DB 3; Length 297;
Best Local Similarity 39.4%; Pred. No. 1.6e-50;
Matches 123; Conservative 61; Mismatches 96; Indels 32; Gaps 8;
QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVAVKKF-VESEDDPVVKKIALREIRMLKQLK 59
Db 1 MEDYTKIEKIGEGTYGVVYKGRHKTTGQVAMKKIRLESEEEG-VPSTAIRISLLKELR 59
QY 60 HNLVNLIEVFRKRKMHVFEYCDHTLLNELELRNPNG--VADGVIKSVLWQTLQALNFC 117
Db 60 HPNIVSLQDVLMOQSRSLYLIFEFLSMDLKKYLDSPPGQYMDSSLVKSYLEQILQGI 119
QY 118 IHNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGD 176
Db 120 HSRRLVLRDLKPQNLLIDDKGTIKLADFLARAFGIPRIVYTHEVVTLWYRSPEVLLG 179
QY 177 YGSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIPIRHQSIFKSN 236
Db 180 RYSTPVDIWSIGTIFAEELATKPKPLFGDSEIDQLFRIFRALG--TPNNE----- 226
QY 237 ISIPEPEMETLEKFSVDVHPVA-----LNFMKGCLKMNPPDDRLTCSQLLESS 284
Db 227 -WPEVESLODYKNTFPKWKPGSLASHVKNLDENGDLDSKMLIYDPAKRISGMALNHP 285
QY 285 YFDSFQEAQIKR 296
Db 286 YFNDL-DNQIKK 296

RESULT 12
US-08-874-347-10
; Sequence 10, Application US/08874347
; Patent No. 5863741
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; TITLE OF INVENTION: CARINII
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/874,347
; FILING DATE: 13-JUN-1997

Fri May 3 11:22:27 2002

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2002, 11:06:49 ; Search time 205.32 Seconds
(without alignments)
538.891 Million cell updates/sec

Title: US-09-671-050-12
Perfect score: 1662
Sequence: 1 MEKYELAKTGEYSYGVFK.....RKARNEGRRRRQQVLP LKS 315

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3516493 seqs, 351254056 residues

Total number of hits satisfying chosen parameters: 3516493

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US080_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/2/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/2/paa/US60_MERGED_COMB.pep1:*
25: /cgn2_6/ptodata/2/paa/US10_MERGED_COMB.pep1:*
26: /cgn2_6/ptodata/2/paa/US09_MERGED_COMB.pep2:*
27: /cgn2_6/ptodata/2/paa/US09_MERGED_COMB.pep1:*
28: /cgn2_6/ptodata/2/paa/US08_MERGED_COMB.pep1:*
29: /cgn2_6/ptodata/2/paa/US07_MERGED_COMB.pep1:*
30: /cgn2_6/ptodata/2/paa/US06_MERGED_COMB.pep1:*
31: /cgn2_6/ptodata/2/paa/PCT_MERGED_COMB.pep1:*
32: /cgn2_6/ptodata/2/paa/PCT_MERGED_COMB.pep1:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1662	100.0	315	20 US-09-671-050-12	Sequence 12, Appli

2	1638	98.6	324	20	US-09-671-050-6	Sequence 6, Appli
3	1636	98.4	347	20	US-09-671-050-10	Sequence 10, Appli
4	1612	97.0	356	20	US-09-671-050-4	Sequence 4, Appli
5	1183.5	71.2	358	18	US-09-411-628-11	Sequence 11, Appli
6	1140	68.6	228	28	US-09-834-496A-2	Sequence 2, Appli
7	1027.5	61.8	392	24	US-60-167-217-16173	Sequence 16173, A
8	1027.5	61.8	392	24	US-60-173-464-13250	Sequence 13250, A
9	1027.5	61.8	392	24	US-60-191-637-16190	Sequence 16190, A
10	1027.5	61.8	392	24	US-60-191-681-12815	Sequence 12815, A
11	1027.5	61.8	392	27	US-09-614-150-16146	Sequence 16146, A
12	986	59.3	290	24	US-60-145-137-509	Sequence 509, App
13	969	58.3	285	24	US-60-146-315-731	Sequence 731, App
14	967.5	58.2	566	18	US-09-411-628-4	Sequence 4, Appli
15	941.5	56.6	493	18	US-09-411-628-10	Sequence 10, Appli
16	791	47.6	187	20	US-09-671-050-2	Sequence 2, Appli
17	791	47.6	198	20	US-09-671-050-8	Sequence 8, Appli
18	641.5	38.6	997	24	US-09-671-050-209	Sequence 209, App
19	581	35.0	305	19	US-09-538-092-1236	Sequence 1236, Ap
20	581	35.0	305	28	US-09-538-092-1236	Sequence 17, Appli
21	581	35.0	333	18	US-09-488-725A-2421	Sequence 2421, Ap
22	574	34.5	352	18	US-09-488-725A-5993	Sequence 5993, Ap
23	570	34.3	298	16	US-09-266-225D-14	Sequence 14, Appli
24	570	34.3	298	17	US-09-338-125-2	Sequence 2, Appli
25	570	34.3	298	21	US-09-771-161A-187	Sequence 187, App
26	569	34.2	191	1	PCT-US01-08656-10259	Sequence 10259, A
27	569	34.2	224	1	PCT-US01-08631-36501	Sequence 36501, A
28	569	34.2	298	18	US-09-411-628-13	Sequence 13, Appli
29	569	34.2	298	28	US-09-953-102A-18	Sequence 18, Appli
30	567	34.1	298	18	US-09-457-040A-29	Sequence 29, Appli
31	567	34.1	298	18	US-09-457-040B-29	Sequence 29, Appli
32	567	34.1	298	19	US-09-538-092-1006	Sequence 1006, Ap
33	567	34.1	298	27	US-09-971-873-29	Sequence 29, Appli
34	567	34.1	298	28	US-09-574-559A-5	Sequence 5, Appli
35	567	34.1	298	32	PCT-US01-30732-4493	Sequence 4493, Ap
36	566.5	34.1	294	24	US-60-324-109-29536	Sequence 29536, A
37	566.5	34.1	294	27	US-09-708-427-58826	Sequence 58826, A
38	565	34.0	286	7	US-08-354-681-11	Sequence 11, Appli
39	556	33.5	297	18	US-09-411-628-12	Sequence 12, Appli
40	556	33.5	297	28	US-09-953-102A-19	Sequence 19, Appli
41	556	33.5	299	27	US-09-760-446A-1544	Sequence 1544, Ap
42	555	33.4	270	3	US-07-857-224A-31	Sequence 31, Appli
43	552.5	33.2	339	18	US-09-464-523-4	Sequence 4, Appli
44	551	33.2	274	1	PCT-US94-11053A-20	Sequence 20, Appli
45	546.5	32.9	297	24	US-60-167-217-11204	Sequence 11204, A

ALIGNMENTS

RESULT 1
US-09-671-050-12
; Sequence 12, Application US/09671050
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: Novel Human Kinase Proteins and
; FILE REFERENCE: LEX-0046-USA
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: US 60/156,511
; PRIOR FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 315
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-671-050-12

Query Match 100.0%; Score 1662; DB 20; Length 315;
Best Local Similarity 100.0%; Pred. No. 5.3e-142;
Matches 315; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKYEKLAKTGE GSYGVVFKCRNKTSGQVVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEKYEKLAKTGE GSYGVVFKCRNKTSGQVVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60

QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEARNPGVADGVKSVLWQTLQALNFCIH 120
Db 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEARNPGVADGVKSVLWQTLQALNFCIH 120

QY 121 NCIHRIKPENILITKOGIIKICDFGFAQIILIPGDAYTDYVATRWYRAPELLVGDTOYGS 180
Db 121 NCIHRIKPENILITKOGIIKICDFGFAQIILIPGDAYTDYVATRWYRAPELLVGDTOYGS 180

QY 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLI PRHQSIKSN GFFHGISIP 240
Db 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLI PRHQSIKSN GFFHGISIP 240

QY 241 EPEDMETLEEFSDVHPVALNFMKGCLKMPDDRLTCSQLESSEYFDSFQEAQIKRKARN 300
Db 241 EPEDMETLEEFSDVHPVALNFMKGCLKMPDDRLTCSQLESSEYFDSFQEAQIKRKARN 300

QY 301 EGRNRRRQOVLPLKS 315
Db 301 EGRNRRRQOVLPLKS 315

RESULT 2
US-09-671-050-6
; Sequence 6, Application US/09671050
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: Novel Human Kinase Proteins and
; TITLE OF INVENTION: Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0046-USA
; CURRENT APPLICATION NUMBER: US/09/671,050
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: US 60/156,511
; PRIOR FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 324
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-671-050-6

Query Match 98.6%; Score 1638; DB 20; Length 324;
Best Local Similarity 100.0%; Pred. No. 8.3e-140;
Matches 310; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKYEKLAKTGE GSYGVVFKCRNKTSGQVVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEKYEKLAKTGE GSYGVVFKCRNKTSGQVVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60

QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEARNPGVADGVKSVLWQTLQALNFCIH 120
Db 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEARNPGVADGVKSVLWQTLQALNFCIH 120

QY 121 NCIHRIKPENILITKOGIIKICDFGFAQIILIPGDAYTDYVATRWYRAPELLVGDTOYGS 180
Db 121 NCIHRIKPENILITKOGIIKICDFGFAQIILIPGDAYTDYVATRWYRAPELLVGDTOYGS 180

QY 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLI PRHQSIKSN GFFHGISIP 240
Db 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLI PRHQSIKSN GFFHGISIP 240

QY 241 EPEDMETLEEFSDVHPVALNFMKGCLKMPDDRLTCSQLESSEYFDSFQEAQIKRKARN 300
Db 241 EPEDMETLEEFSDVHPVALNFMKGCLKMPDDRLTCSQLESSEYFDSFQEAQIKRKARN 300

QY 301 EGRNRRRQO V 310
Db 301 EGRNRRRQO V 310

RESULT 3
US-09-671-050-10
; Sequence 10, Application US/09671050
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: Novel Human Kinase Proteins and
; TITLE OF INVENTION: Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0046-USA
; CURRENT APPLICATION NUMBER: US/09/671,050
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: US 60/156,511
; PRIOR FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 347
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-671-050-10

Query Match 98.4%; Score 1636; DB 20; Length 347;
Best Local Similarity 90.8%; Pred. No. 1.4e-139;
Matches 315; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

QY 1 MEKYEKLAKTGE GSYGVVFKCRNKTSGQVVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEKYEKLAKTGE GSYGVVFKCRNKTSGQVVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60

QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEARNPGVADGVKSVLWQTLQALNFCIH 120
Db 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEARNPGVADGVKSVLWQTLQALNFCIH 120

QY 121 NCIHRIKPENILITKOGIIKICDFGFAQIILIPGDAYTDYVATRWYRAPELLVGDTOYGS 180
Db 121 NCIHRIKPENILITKOGIIKICDFGFAQIILIPGDAYTDYVATRWYRAPELLVGDTOYGS 180

QY 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTL----- 217
Db 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLVETGFRHVDQAGLELLTSSDPPA 240

QY 218 -----GKLI PRHQSIKSN GFFHGISIPEDMETLEEFSDVHPVALNFMKGCLK 268
Db 241 VASQSAGITGKLI PRHQSIKSN GFFHGISIPEDMETLEEFSDVHPVALNFMKGCLK 300

QY 269 MNPDDRLTCSQLESSEYFDSFQEAQIKRKARNRNRROQVLP LKS 315
Db 301 MNPDDRLTCSQLESSEYFDSFQEAQIKRKARNRNRROQVLP LKS 347

RESULT 4
US-09-671-050-4
; Sequence 4, Application US/09671050
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory

```

; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: Novel Human Kinase Proteins and
; FILE REFERENCE: LEX-0046-USA
; CURRENT APPLICATION NUMBER: US/09/671,050
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: US 60/156,511
; PRIOR FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 356
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-671-050-4

Query Match      97.0%; Score 1612; DB 20; Length 356;
Best Local Similarity 90.6%; Pred. No. 2.2e-137;
Matches 310; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

QY 1 MEKYEKLAKTGECSYGVVFKCRNKTSQGVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEKYEKLAKTGECSYGVVFKCRNKTSQGVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
QY 61 PNLVNLIEVFRRRKRMHLVFEYCDHTLLNELERNPNGVADGVIKSVLWQTLQALNFCIH 120
Db 61 PNLVNLIEVFRRRKRMHLVFEYCDHTLLNELERNPNGVADGVIKSVLWQTLQALNFCIH 120
QY 121 NCIHRIKPKENILITKQGIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDTOYGS 180
Db 121 NCIHRIKPKENILITKQGIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDTOYGS 180
QY 121 NCIHRIKPKENILITKQGIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDTOYGS 180
Db 121 NCIHRIKPKENILITKQGIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDTOYGS 180
QY 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTL----- 217
Db 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLVETGFRHVDQAGLELLTSSDP 240
QY 218 -----GKLIPRHQSIKSNGGFFHGISIPEDMETLEEKFSVHPVALNFMKGCLK 268
Db 241 VASQSAGITGKLIPRHQSIKSNGGFFHGISIPEDMETLEEKFSVHPVALNFMKGCLK 300
QY 269 MNPDDLRLTCSQLLESSYFDSFQEAQIKRKARNEGRNRRQOV 310
Db 301 MNPDDLRLTCSQLLESSYFDSFQEAQIKRKARNEGRNRRQOV 342

RESULT 5
US-09-411-628-11
; Sequence 11, Application US/09411628
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: cDNA, GENOMIC, AND PREDICTED PROTEIN
; TITLE OF INVENTION: SEQUENCES OF LEARNING-INDUCED KINASES
; FILE REFERENCE: 13761-707
; CURRENT APPLICATION NUMBER: US/09/411,628
; CURRENT FILING DATE: 1999-10-01
; EARLIER APPLICATION NUMBER: US 60/102,906
; EARLIER FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-411-628-11

Query Match      71.2%; Score 1183.5; DB 18; Length 358;
Best Local Similarity 71.8%; Pred. No. 1.7e-98;
```

```

Matches 211; Conservative 41; Mismatches 41; Indels 1; Gaps 1;

QY 1 MEKYEKLAKTGECSYGVVFKCRNKTSQGVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 2 MEKYEKIGKIGESYGVVFKCRNRDTQIVAIKKFLESEDDPVVKKIALREIRMLKQLKH 61
QY 61 PNLVNLIEVFRRRKRMHLVFEYCDHTLLNELERNPNGVADGVIKSVLWQTLQALNFCIH 120
Db 62 PNLVNLIEVFRRRKRMHLVFEYCDHTVLHELDYRQGVPEHLVKSITWQTLQAVNFCCHK 121
QY 121 NCIHRIKPKENILITKQGIKICDFGFAQILI-PGDAYTDYVATRWYRAPELLVGDTOYG 179
Db 122 NCIHRIKPKENILITKHSVIKLCDFGARLLTGSDYITDYVATRWYRSPELLVGDTOYG 181
QY 180 SSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKIIPRHQSIKSNFPFHGISI 239
Db 182 PPDVVAIGCVFAELLTSGVPLWPGKSDVDQLYLIIRKTLGDLIPRHQQVFSTNQYFSGVKI 241
QY 240 PEPEDMETLEEKFSVHPVALNFMKGKGLKMNPDRLTCSQLLESSYFDSFQEAQ 293
Db 242 PDPEDMEPLELKFPNISYPALGILLKGLHMDPTERTLTCEQLLHHPYFENIREIE 295

RESULT 6
US-09-834-496A-2
; Sequence 2, Application US/09834496A
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann
; TITLE OF INVENTION: 14257 NOVEL PROTEIN KINASE MOLECULES AND
; TITLE OF INVENTION: THEIR USES THEREFOR
; FILE REFERENCE: 381552000900
; CURRENT APPLICATION NUMBER: US/09/834,496A
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/196,910
; PRIOR FILING DATE: 2000-04-13
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-496A-2

Query Match      68.6%; Score 1140; DB 28; Length 228;
Best Local Similarity 98.2%; Pred. No. 8.4e-95;
Matches 214; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MEKYEKLAKTGECSYGVVFKCRNKTSQGVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MGKYEKLAKTGECSYGVVFKCRNKTSQGVAVKKFVESEDDPIVKKIALREIRMLKQLKH 60
QY 61 PNLVNLIEVFRRRKRMHLVFEYCDHTLLNELERNPNGVADGVIKSVLWQTLQALNFCIH 120
Db 61 PNLVNLIEVFRRRKRMHLVFEYCDHALLNELERNPNGVADGVIKSVLWQTLQALNFCIH 120
QY 121 NCIHRIKPKENILITKQGIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDTOYGS 180
Db 121 NCIHRIKPKENILITKQGIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDTOYGS 180
QY 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLG 218
Db 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLG 218

RESULT 7
US-60-167-217-16173
; Sequence 16173, Application US/60167217
; GENERAL INFORMATION:
; APPLICANT: Li, Peter W. D.
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING DROSOPHILA PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
```

```
; FILE REFERENCE: CL000152
; CURRENT APPLICATION NUMBER: US/60/167,217
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 23195
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16173
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Drosophila
US-60-167-217-16173

Query Match      61.8%; Score 1027.5; DB 24; Length 392;
Best Local Similarity 58.4%; Pred. No. 2.8e-84;
Matches 184; Conservative 65; Mismatches 61; Indels 5; Gaps 2;

Qy 1 MEKYEKLAKTGE GSYGVVFKCRNKTSQGVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MDRYEKLSRLGEGSYGVVYKCRDRGTGALVAVKRFVESEDDPAIRKIALREIRLLKNLKH 60

Qy 61 PNLVNLIEVFRKKRMHLVFEYCDHTLLNELEARNPNGVADGVIKSVLWQTLOALNFCCHI 120
Db 61 PNLVSLLEVFRRKRRRLHLVFECELTVLHELEHRHQGCPPEHLTKQICYQTLGVAYCHKQ 120

Qy 121 NCIHRODKPENILITKQGIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDYQYGS 180
Db 121 GCLHRDIKPENILLTAQGVKLCDFGFARMLSPGENYTDYVATRWYRAPELLVGDYQYGT 180

Qy 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIPRHQSIFKSNGFHGISIP 240
Db 181 PVDVWAIGCLFAELVRGEALWPGRSVDQLYLIRKTLGDLPLRHQIFGQNEYFKGITLP 240

Qy 241 EPEDMETLEEKF---SDVHPVALNFMKGCLKMNPDDRLTCSQLLESSYFDSF--QEAQIK 295
Db 241 VPPTLEPLEDKMPAKSQQNPLTIDFLKKCLDKDPTKRWSCCKLTKHSYFDDYIAKQRELE 300

Qy 296 RKARNEGRNRRRQV 310
Db 301 HVNSLEAANLRQQQL 315

Query Match      61.8%; Score 1027.5; DB 24; Length 392;
Best Local Similarity 58.4%; Pred. No. 2.8e-84;
Matches 184; Conservative 65; Mismatches 61; Indels 5; Gaps 2;

Qy 1 MEKYEKLAKTGE GSYGVVFKCRNKTSQGVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MDRYEKLSRLGEGSYGVVYKCRDRGTGALVAVKRFVESEDDPAIRKIALREIRLLKNLKH 60

Qy 61 PNLVNLIEVFRKKRMHLVFEYCDHTLLNELEARNPNGVADGVIKSVLWQTLOALNFCCHI 120
Db 61 PNLVSLLEVFRRKRRRLHLVFECELTVLHELEHRHQGCPPEHLTKQICYQTLGVAYCHKQ 120

Qy 121 NCIHRODKPENILITKQGIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDYQYGS 180
Db 121 GCLHRDIKPENILLTAQGVKLCDFGFARMLSPGENYTDYVATRWYRAPELLVGDYQYGT 180

Qy 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIPRHQSIFKSNGFHGISIP 240
Db 181 PVDVWAIGCLFAELVRGEALWPGRSVDQLYLIRKTLGDLPLRHQIFGQNEYFKGITLP 240

Qy 241 EPEDMETLEEKF---SDVHPVALNFMKGCLKMNPDDRLTCSQLLESSYFDSF--QEAQIK 295
Db 241 VPPTLEPLEDKMPAKSQQNPLTIDFLKKCLDKDPTKRWSCCKLTKHSYFDDYIAKQRELE 300

Qy 296 RKARNEGRNRRRQV 310
Db 301 HVNSLEAANLRQQQL 315

Query Match      61.8%; Score 1027.5; DB 24; Length 392;
Best Local Similarity 58.4%; Pred. No. 2.8e-84;
Matches 184; Conservative 65; Mismatches 61; Indels 5; Gaps 2;

Qy 1 MEKYEKLAKTGE GSYGVVFKCRNKTSQGVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MDRYEKLSRLGEGSYGVVYKCRDRGTGALVAVKRFVESEDDPAIRKIALREIRLLKNLKH 60

Qy 61 PNLVNLIEVFRKKRMHLVFEYCDHTLLNELEARNPNGVADGVIKSVLWQTLOALNFCCHI 120
Db 61 PNLVSLLEVFRRKRRRLHLVFECELTVLHELEHRHQGCPPEHLTKQICYQTLGVAYCHKQ 120

Qy 121 NCIHRODKPENILITKQGIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDYQYGS 180
Db 121 GCLHRDIKPENILLTAQGVKLCDFGFARMLSPGENYTDYVATRWYRAPELLVGDYQYGT 180

Qy 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIPRHQSIFKSNGFHGISIP 240
Db 181 PVDVWAIGCLFAELVRGEALWPGRSVDQLYLIRKTLGDLPLRHQIFGQNEYFKGITLP 240

Qy 241 EPEDMETLEEKF---SDVHPVALNFMKGCLKMNPDDRLTCSQLLESSYFDSF--QEAQIK 295
Db 241 VPPTLEPLEDKMPAKSQQNPLTIDFLKKCLDKDPTKRWSCCKLTKHSYFDDYIAKQRELE 300

Qy 296 RKARNEGRNRRRQV 310
Db 301 HVNSLEAANLRQQQL 315

Query Match      61.8%; Score 1027.5; DB 24; Length 392;
Best Local Similarity 58.4%; Pred. No. 2.8e-84;
Matches 184; Conservative 65; Mismatches 61; Indels 5; Gaps 2;

Qy 1 MEKYEKLAKTGE GSYGVVFKCRNKTSQGVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MDRYEKLSRLGEGSYGVVYKCRDRGTGALVAVKRFVESEDDPAIRKIALREIRLLKNLKH 60

Qy 61 PNLVNLIEVFRKKRMHLVFEYCDHTLLNELEARNPNGVADGVIKSVLWQTLOALNFCCHI 120
Db 61 PNLVSLLEVFRRKRRRLHLVFECELTVLHELEHRHQGCPPEHLTKQICYQTLGVAYCHKQ 120

Qy 121 NCIHRODKPENILITKQGIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDYQYGS 180
Db 121 GCLHRDIKPENILLTAQGVKLCDFGFARMLSPGENYTDYVATRWYRAPELLVGDYQYGT 180

Qy 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIPRHQSIFKSNGFHGISIP 240
Db 181 PVDVWAIGCLFAELVRGEALWPGRSVDQLYLIRKTLGDLPLRHQIFGQNEYFKGITLP 240

Qy 241 EPEDMETLEEKF---SDVHPVALNFMKGCLKMNPDDRLTCSQLLESSYFDSF--QEAQIK 295
Db 241 VPPTLEPLEDKMPAKSQQNPLTIDFLKKCLDKDPTKRWSCCKLTKHSYFDDYIAKQRELE 300

Qy 296 RKARNEGRNRRRQV 310
Db 301 HVNSLEAANLRQQQL 315
```

```
Qy 121 NCIHRODKPENILITKQGIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDYQYGS 180
Db 121 GCLHRDIKPENILLTAQGVKLCDFGFARMLSPGENYTDYVATRWYRAPELLVGDYQYGT 180

Qy 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIPRHQSIFKSNGFHGISIP 240
Db 181 PVDVWAIGCLFAELVRGEALWPGRSVDQLYLIRKTLGDLPLRHQIFGQNEYFKGITLP 240

Qy 241 EPEDMETLEEKF---SDVHPVALNFMKGCLKMNPDDRLTCSQLLESSYFDSF--QEAQIK 295
Db 241 VPPTLEPLEDKMPAKSQQNPLTIDFLKKCLDKDPTKRWSCCKLTKHSYFDDYIAKQRELE 300

Qy 296 RKARNEGRNRRRQV 310
Db 301 HVNSLEAANLRQQQL 315

RESULT 9
US-60-191-637-16190
; Sequence 16190, Application US/60191637
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: NUCLEIC ACID DETECTION KITS COMPRISING
; TITLE OF INVENTION: GENE SEQUENCES EXPRESSED FROM THE DROSOPHILA GENOME, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000392
; CURRENT APPLICATION NUMBER: US/60/191,637
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 42660
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16190
; LENGTH: 392
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-60-191-637-16190

Query Match      61.8%; Score 1027.5; DB 24; Length 392;
Best Local Similarity 58.4%; Pred. No. 2.8e-84;
Matches 184; Conservative 65; Mismatches 61; Indels 5; Gaps 2;

Qy 1 MEKYEKLAKTGE GSYGVVFKCRNKTSQGVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MDRYEKLSRLGEGSYGVVYKCRDRGTGALVAVKRFVESEDDPAIRKIALREIRLLKNLKH 60

Qy 61 PNLVNLIEVFRKKRMHLVFEYCDHTLLNELEARNPNGVADGVIKSVLWQTLOALNFCCHI 120
Db 61 PNLVSLLEVFRRKRRRLHLVFECELTVLHELEHRHQGCPPEHLTKQICYQTLGVAYCHKQ 120

Qy 121 NCIHRODKPENILITKQGIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDYQYGS 180
Db 121 GCLHRDIKPENILLTAQGVKLCDFGFARMLSPGENYTDYVATRWYRAPELLVGDYQYGT 180

Qy 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIPRHQSIFKSNGFHGISIP 240
Db 181 PVDVWAIGCLFAELVRGEALWPGRSVDQLYLIRKTLGDLPLRHQIFGQNEYFKGITLP 240

Qy 241 EPEDMETLEEKF---SDVHPVALNFMKGCLKMNPDDRLTCSQLLESSYFDSF--QEAQIK 295
Db 241 VPPTLEPLEDKMPAKSQQNPLTIDFLKKCLDKDPTKRWSCCKLTKHSYFDDYIAKQRELE 300

Qy 296 RKARNEGRNRRRQV 310
Db 301 HVNSLEAANLRQQQL 315

RESULT 10
US-60-191-681-12815
; Sequence 12815, Application US/60191681
; GENERAL INFORMATION:
; APPLICANT: Li, Peter, W.D.
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING ISOLATED DROSOPHILA PROTEINS AND
```


QY 181 SVDIWAIGCVFAELLTQGPLWPKSDVDQLYLIIIRTLGKLIIPRHSIFKNGFFHGISIP 240
 ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||
 Db 180 PVDVWAIGCFLAELVRGEALWPGRSDVDQLYLIRKLVLGDLLPRHIQIFGONEYFKGITLP 239

```

Qy      241 EPEDMETLEKF---SDVHPVALNFMKGCLKNPDDRLTC$QLLESSYFD 287
        | : ||| | | ::::| : | : | : | : | : | : | : | : | : | : |
Db      240 VPPTLEPLEDKMPAKSQQNPLTIDFKKXXKDPTKRWCSEKITKHSYFD 289

```

RESULT 13

US-60-146-315-731

```

RESULT      13
US-60-146-315-731
; Sequence 731, Application US/60146315
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED KINASE PROTEIN
; TITLE OF INVENTION: MOLECULES ENCODING KINASE
; FILE REFERENCE: CL000064
; CURRENT APPLICATION NUMBER: US/60/146.315
; CURRENT FILING DATE: 1999-07-29
; NUMBER OF SEQ ID NOS: 1086
; SOFTWARE: FastSEQ for Windows Version 3.0.0
; SEQ ID NO 731
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Drosophila
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)..(285)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-146-315-731

```

Query Match	58.3%	Score	969;	DB 24;	Length	285;
Best Local Similarity	60.8%;	Pred. No.	3.8e-79;			
Matches`174;	Conservative	55;	Mismatches	53;	Indels	4;
					Gaps	2;

Qy	4	YKLAKTGEGSVGVFKCRNKTSGVVAVKRFESEDDPVVKKIALREIRMLKQLKHPNL	63
		: : : : : : : :	
Db	1	YEKLSRLGEGSVGVVKCRDRETGALVAVKRFESEDDPAIRKIALREIRLLK-LKHPNL	59

Qy	64	VNLEVFRRKRKMHLVFEYCDHTLNELEERNPGVADGVIKSVLTWOTLQALNPFCHIHNCI	123
		: : : : : : : : : : : : : : : : : : :	
Db	60	VSLEVFRRKRLHLVXFECFLTVLHELERHPGCGPEHITKQICQTTLXGVAYCHKQGCL	119

```
Qy 124 HRDIKPENILITKOGIIKICDFGAQILIPGDAYTDYVATRWYRAPELLVGDTQYGSSVD 183
    | ||||| : | : ||||| : || : ||||| ||||| : ||
Db 120 HXDIKPENILLTAQGQVKLCDFGFARMLSPGENYTDYVATRWYRAPELLVGDTQYGTVPD 179
```

[illegible]

Qy	244	DMETLEKFT--SDVHPVALNFMKGLKMPNDDRLTCSQLLESSYF	2866
		: : : : : : : : : : : : :	
Db	240	TLEPLEDKMPAKSQONPLTIDFLKCCXXKDPTRWCCEKLTHTSYF	2855

```

RESULT 14
US-09-411-628-4
; Sequence 4, Application US/09411628
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
; TITLE OF INVENTION: SEQUENCES OF LEARNING-INDUCED KINASES
; FILE REFERENCE: 13761-707
; CURRENT APPLICATION NUMBER: US/09/411.628
; CURRENT FILING DATE: 1999-10-01
; EARLIER APPLICATION NUMBER: US 60/102.906
; EARLIER FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4

```

```
; LENGTH: 566
; TYPE: PRT
; ORGANISM: Orcyotolagus cuniculus
US-09-411-628-4

Query Match          58.2%; Score 967.5; DB 18; Length 566;
Best Local Similarity 54.6%; Pred. No. 1.3e-78;
Matches 177; Conservative 67; Mismatches 65; Indels 15; Gaps

QY      1 MEKYEKLAKTGEGSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
        ||||| | ||||| : |||| |:::||:||||: |||||:||||:||||:|
Db      1 MEKYENLGLVGESYGMYMKCRNKDSGRIVAIKKFLESDDDKMVKKIAMREIKLLKQLRH 60
        ||||| | ||||| : |||| |:::||:||||: |||||:||||:||||:|

QY      61 PNLVNLIEVRKRKRMHLVFECYCDHTLLNELRNPNGVADGVIKSVLWQTLOALNFCHIH 120
        |||||:| ::::: :|||: |||:|:::|| |||: | ::: | ||| |
Db      61 ENLVNLLEVCKKKRWYLVFVVDHTILDDLELPNGLDQVVQKYLFQIINGIGFCHSH 120
        |||||:| ::::: :|||: |||:|:::|| |||: | ::: | ||| |

QY      121 NCIHRDIKPENILITKGIIKICDFGAQIL-IPGDAYTDYVATRWYRAPELLVGDTQYG 179
        | ||||| |||||: : (: : ||||| : | ||: ||||| ||||| ||||| : ||
Db      121 NIHRDIKPENILVSQGVVKLCDFGFARTLAAPCEVYTDYVATRWYRAPELLVGDKYKG 180
        | ||||| |||||: : (: : ||||| : | ||: ||||| ||||| ||||| : ||

QY      180 SSVDIWAIGCVFAELLTQCPLWPGGSDVDQLYLIIRTCLKLIPRHSIFKSNGFFHGISI 239
        : ||: ||||: | : | : ||: || |:::|||: || |||||: | | | : | :
Db      181 KAVDWAIGCLVTETLMGEPLFPGSDSIDQLVLIMRCGLNIIPRHQELFYKNPVPFAGVRL 240
        : ||: ||||: | : | : ||: || |:::|||: || |||||: | | | : | :

QY      240 PEPEDMETLEKFSDVHPVALNFMKGCLKMNPDDRLLTCSSOLLESSYF--DSF-----QEA 292
        || : || : || : : | : | : || : || |:::|||: || |||||: | | | : ||
Db      241 PEIKESEPLERRYPKLSEVVIDLAKKCLHVDPKRPFCAEILLHHDDFFQMCGFAERFSQEL 300
        || : || : || : : | : | : || : || |:::|||: || |||||: | | | : ||

QY      293 QIK--RKARN-----EGRNRRRQQ 309
        (: : | : ||) : || : : :
Db      301 QMKVQKDARNISLSKKKSQRKKKEK 324
```

```

RESULT 15
US-09-411-628-10
; Sequence 10, Application US/09411628
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
; TITLE OF INVENTION: SEQUENCES OF LEARNING-INDUCED KINASES
; FILE REFERENCE: 13761-707
; CURRENT APPLICATION NUMBER: US/09/411,628
; CURRENT FILING DATE: 1999-10-01
; EARLIER APPLICATION NUMBER: US 60/102,906
; EARLIER FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-411-628-10

```

Query Match	56.6%;	Score 941.5;	DB 18;	Length 493;
Best Local Similarity	53.1%;	Pred. No. 2.5e-76;		
Matches 172;	Conservative 68;	Mismatches 69;	Indels 15;	Gaps 5;

	Qy	1	MEKYEKLAKTGEGSYGVVFKCRNKTSGQVVAVKKFVESEDDPVVKKTALRETRMLKQLKH	60
	Dd	1	MEKENLGLVGEGSYGMVMKCRNKDTGRIVATKFFLESDDDKMKVKTIAMREIKLLKQLRH	60
	Qy	61	PNLVNLIJEFRKRKMHLVFYCDHTLLNELERNPNGVADGVIKSWLMQTLOALNFCHTH	120
	Dd	61	ENLVNLEVCRRKKRWYLFVEFVDHITLDDELFPNGLDYQQWVKYLEFIINGIGFCHSH	120
	Qy	121	NCTHRDIKPENILJTQGIIKICDFGAQILL-IPGDAYTDYVATRMWRAPELLVGDTQYG	179
	Dd	121	NIITHRDIKPENILVSQSQVVKLCDFGFARTLAAPGEVYTDYVATRMWRAPELLVGDKYKG	180
	Qy	180	SSVDIWAIGCVFAELLTGQPLWPBGKSDDVDQLYIIRTLGKLIPRHQSIFKSNGFHFGISI	239

us-09-671-050-12.rapm

Fri May 3 11:22:27 2002

Db 181 KAVDVWAIGCLVTEMFMGEPLFPGDSIDQLYHIMMCLGNLIPRHQELFNKNPVPFAGVRL 240
Qy 240 PEPEDMETLEEKFSVDVHPVALNFMKGCLKMNPDRLTCSQLESSEYF--DSF-----QEA 292
Db 241 PEIKEREPLERRYKLSSEVIDLAKKCLHIDPDKRPFCALLHHDFFQMDGFAERFSQEL 300
Qy 293 QIK--RKARN-----EGRNRRROQ 309
Db 301 QLKVQKDARNVSLSKKSONRKKEK 324

Search completed: May 3, 2002, 11:06:50
Job time: 324 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 17:10:06 ; Search time 1547.66 Seconds
(without alignments)
11096.471 Million cell updates/sec

Title: US-09-671-050-9
Perfect score: 1041
Sequence: 1 atggaaaagtatgaaaaatt.....aggttacttcgcgtcaaaagt 1041

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1472140 seqs, 8248589755 residues

Word size : 20

Total number of hits satisfying chosen parameters: 35700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

GenEmbl:*

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_om:*
- 20: em_or:*
- 21: em_ov:*
- 22: em_pat:*
- 23: em_ph:*
- 24: em_pl:*
- 25: em_ro:*
- 26: em_sts:*
- 27: em_sy:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htgo_hum:*
- 31: em_htgo_inv:*
- 32: em_htgo_rod:*
- 33: em_htg_hum:*
- 34: em_htg_inv:*
- 35: em_htg_rod:*
- 36: em_htg_other:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

*

Result No.	Score	Query Match	Length	DB	ID	Description
1	1041	100.0	1041	6	AX107720	AX107720 Sequence
2	1025	98.5	1068	6	AX107714	AX107714 Sequence
3	652	62.6	945	6	AX107722	AX107722 Sequence
4	652	62.6	972	6	AX107716	AX107716 Sequence
5	572	54.9	1819	6	AX107724	AX107724 Sequence
6	454	43.6	561	6	AX107712	AX107712 Sequence
7	454	43.6	594	6	AX107718	AX107718 Sequence
8	454	43.6	1083	6	AX166534	AX166534 Sequence
9	368	35.4	911	6	AX056404	AX056404 Sequence
10	198	19.0	80362	9	AC079615	AC079615 Homo sapi
11	198	19.0	155028	33	AC037489	AC037489 Homo sapi
12	198	19.0	164281	2	AC092672	AC092672 Homo sapi
13	59	5.7	155028	33	AC037489	AC037489 Homo sapi
14	47	4.5	10135	9	GGO312524	AJ312524 Gorilla g
15	47	4.5	10135	9	GGO312526	AJ312526 Gorilla g
16	47	4.5	10135	9	GGO312527	AJ312527 Gorilla g
17	47	4.5	10135	9	GGO312528	AJ312528 Gorilla g
18	47	4.5	10139	9	GGO241093	AJ241093 Gorilla g
19	47	4.5	10139	9	GGO312522	AJ312522 Gorilla g
20	47	4.5	10139	9	GGO312529	AJ312529 Gorilla g
21	47	4.5	10139	9	GGO312530	AJ312530 Gorilla g
22	47	4.5	10141	9	GGO312521	AJ312521 Gorilla g
23	47	4.5	10142	9	GGO312523	AJ312523 Gorilla g
24	47	4.5	10142	9	GGO312525	AJ312525 Gorilla g
25	47	4.5	10145	9	PPA270090	AJ270090 Pan panis
26	47	4.5	10145	9	PPA270091	AJ270091 Pan panis
27	47	4.5	10145	9	PPA270093	AJ270093 Pan panis
28	47	4.5	10145	9	PPA270094	AJ270094 Pan panis
29	47	4.5	10151	9	PPA270092	AJ270092 Pan panis
30	47	4.5	10151	9	PTR270065	AJ270065 Pan trogl
31	47	4.5	10151	9	PTR270066	AJ270066 Pan trogl
32	47	4.5	10151	9	PTR270068	AJ270068 Pan trogl
33	47	4.5	10151	9	PTR270070	AJ270070 Pan trogl
34	47	4.5	10151	9	PTR270071	AJ270071 Pan trogl
35	47	4.5	10151	9	PTR270072	AJ270072 Pan trogl
36	47	4.5	10152	9	PTR241092	AJ241092 Pan trogl
37	47	4.5	10154	9	PTR270061	AJ270061 Pan trogl
38	47	4.5	10154	9	PTR270062	AJ270062 Pan trogl
39	47	4.5	10154	9	PTR270064	AJ270064 Pan trogl
40	47	4.5	10154	9	PTR270067	AJ270067 Pan trogl
41	47	4.5	10154	9	PTR270073	AJ270073 Pan trogl
42	47	4.5	10154	9	PTR270076	AJ270076 Pan trogl
43	47	4.5	10154	9	PTR270077	AJ270077 Pan trogl
44	47	4.5	10154	9	PTR270078	AJ270078 Pan trogl
45	47	4.5	10154	9	PTR270079	AJ270079 Pan trogl
46	47	4.5	10154	9	PTR270080	AJ270080 Pan trogl
47	47	4.5	10154	9	PTR270081	AJ270081 Pan trogl
48	47	4.5	10154	9	PTR270082	AJ270082 Pan trogl
49	47	4.5	10154	9	PTR270083	AJ270083 Pan trogl
50	47	4.5	10154	9	PTR270084	AJ270084 Pan trogl
51	47	4.5	10154	9	PTR270085	AJ270085 Pan trogl
52	47	4.5	10154	9	PTR270086	AJ270086 Pan trogl
53	47	4.5	10154	9	PTR270087	AJ270087 Pan trogl
54	47	4.5	10154	9	PTR270088	AJ270088 Pan trogl
55	47	4.5	10154	9	PTR270089	AJ270089 Pan trogl
56	47	4.5	10155	9	PTR270075	AJ270075 Pan trogl
57	47	4.5	10156	9	PTR270074	AJ270074 Pan trogl
58	47	4.5	61359	9	AC079396	AC079396 Homo sapi
59	47	4.5	64753	9	AL355302	AL355302 Human DNA
60	47	4.5	80668	9	AC005606	AC005606 Homo sapi
61	47	4.5	99411	9	AC005207	AC005207 Homo sapi
62	47	4.5	118826	2	AC092473	AC092473 Homo sapi
63	47	4.5	124821	9	AC004983	AC004983 Homo sapi
64	47	4.5	125673	9	AL356796	AL356796 Human DNA
65	47	4.5	126543	2	AC026156	AC026156 Homo sapi
66	47	4.5	127683	9	AC003683	AC003683 Homo sapi
67	47	4.5	132292	9	AL138895	AL138895 Human DNA
68	47	4.5	145673	2	AL391068	AL391068 Homo sapi
69	47	4.5	153405	2	AL357252	AL357252 Homo sapi
70	47	4.5	153452	2	AC044895	AC044895 Homo sapi

71	47	4.5	156284	2	AC036154	AC036154	Homo sapi	c 144	41	3.9	174442	9	CNS05TCW	AL355885	Human chr
72	47	4.5	156294	2	AC092409	AC092409	Papio cyn	c 145	41	3.9	177352	2	AL355338	AL355338	Homo sapi
c 73	47	4.5	160156	9	AC009362	AC009362	Homo sapi	146	41	3.9	184737	2	AC087833	AC087833	Papio cyn
c 74	47	4.5	163898	2	AC025255	AC025255	Homo sapi	c 147	41	3.9	188079	2	AL158209	AL158209	Homo sapi
75	47	4.5	169065	2	AP002338	AP002338	Homo sapi	148	41	3.9	190846	9	AC006372	AC006372	Homo sapi
c 76	47	4.5	171627	2	AF336382	AF336382	Homo sapi	149	41	3.9	193390	2	AC073261	AC073261	Homo sapi
77	47	4.5	171801	2	AC080162	AC080162	Homo sapi	c 150	41	3.9	194265	2	AL590644	AL590644	Homo sapi
c 78	47	4.5	172136	2	AC023959	AC023959	Homo sapi	151	41	3.9	197212	2	AC018723	AC018723	Homo sapi
79	47	4.5	172279	2	AC023202	AC023202	Homo sapi	152	41	3.9	219935	9	AC005015	AC005015	Homo sapi
80	47	4.5	172609	2	AP001463	AP001463	Homo sapi	153	41	3.9	285798	2	AL589862	AL589862	Homo sapi
c 81	47	4.5	172609	2	AP001463	AP001463	Homo sapi	c 154	40	3.8	61615	9	AC068779	AC068779	Homo sapi
c 82	47	4.5	172769	2	AC018628	AC018628	Homo sapi	155	40	3.8	73431	2	AP000575	AP000575	Homo sapi
c 83	47	4.5	175684	2	AC073343	AC073343	Homo sapi	c 156	40	3.8	109290	2	HS838114	Y12335	Homo sapien
c 84	47	4.5	177003	2	AC022043	AC022043	Homo sapi	c 157	40	3.8	117621	9	AC003950	AC003950	Homo sapi
c 85	47	4.5	178828	2	AL441925	AL441925	Homo sapi	c 158	40	3.8	118218	2	AP000670	AP000670	Homo sapi
c 86	47	4.5	182387	2	AL390246	AL390246	Homo sapi	159	40	3.8	129272	9	AL451125	AL451125	Human DNA
c 87	47	4.5	186349	2	AC073420	AC073420	Homo sapi	c 160	40	3.8	149597	2	AC034271	AC034271	Homo sapi
88	47	4.5	201470	2	AC073924	AC073924	Homo sapi	161	40	3.8	149758	2	AL591163	AL591163	Homo sapi
c 89	47	4.5	202397	2	AC074394	AC074394	Homo sapi	162	40	3.8	150660	9	AL157372	AL157372	Human DNA
c 90	47	4.5	209855	2	AL592309	AL592309	Homo sapi	163	40	3.8	157454	2	AC079929	AC079929	Homo sapi
c 91	47	4.5	212074	2	AC010321	AC010321	Homo sapi	164	40	3.8	162658	2	AC093081	AC093081	Homo sapi
c 92	47	4.5	216477	2	AC079882	AC079882	Homo sapi	c 165	40	3.8	163577	9	AL590428	AL590428	Human DNA
c 93	44	4.2	201227	2	AC092171	AC092171	Homo sapi	166	40	3.8	167852	2	AP001187	AP001187	Homo sapi
c 94	44	4.2	214930	9	AL136084	AL136084	Human DNA	c 167	40	3.8	169466	9	CNS01DW7	AL136419	Human chr
c 95	44	4.2	172170	2	AC018357	AC018357	Homo sapi	170	40	3.8	169825	9	AC012652	AC012652	Homo sapi
c 96	43	4.2	180129	2	AC021026	AC021026	Homo sapi	171	40	3.8	170200	9	AC008012	AC008012	Homo sapi
c 97	43	4.1	182834	2	AC027523	AC027523	Homo sapi	172	40	3.8	171980	2	AP000928	AP000928	Homo sapi
c 98	43	4.1	194198	2	AC011774	AC011774	Homo sapi	174	40	3.8	178100	2	AP001558	AP001558	Homo sapi
c 99	43	4.1	200582	2	AC008746	AC008746	Homo sapi	175	40	3.8	179859	9	AC007563	AC007563	Homo sapi
c 100	42	4.0	145465	2	AC008894	AC008894	Homo sapi	c 176	40	3.8	180068	2	AC074213	AC074213	Homo sapi
c 101	42	4.0	49201	2	AC009100	AC009100	Homo sapi	c 177	40	3.8	181975	2	AC079406	AC079406	Homo sapi
c 102	42	4.0	49201	2	AC009100	AC009100	Homo sapi	178	40	3.8	185820	9	HS1009E24	AL109804	Human DNA
c 103	42	4.0	175046	2	AC010531	AC010531	Homo sapi	179	40	3.8	187483	2	AC090316	AC090316	Homo sapi
c 104	42	4.0	179034	33	AC009109	AC009109	Homo sapi	c 180	40	3.8	187939	2	AC048338	AC048338	Homo sapi
c 105	42	4.0	180315	2	AC068460	AC068460	Homo sapi	c 181	40	3.8	190739	2	AP001361	AP001361	Homo sapi
c 106	42	4.0	181609	9	AC079385	AC079385	Homo sapi	c 182	40	3.8	192298	9	AP000943	AP000943	Homo sapi
c 107	42	4.0	183332	2	AC025521	AC025521	Homo sapi	c 183	40	3.8	195334	2	AC091082	AC091082	Homo sapi
c 108	42	4.0	183332	2	AC025521	AC025521	Homo sapi	184	40	3.8	200026	2	AC090368	AC090368	Homo sapi
c 109	42	4.0	200000	2	AC006294	AC006294	Homo sapi	c 185	40	3.8	201508	2	AC026290	AC026290	Homo sapi
c 110	42	4.0	175046	2	AC010531	AC010531	Homo sapi	c 186	40	3.8	203300	9	AC000134	AC000134	Homo sapi
c 111	42	4.0	175046	2	AC010531	AC010531	Homo sapi	187	39	3.7	10163	9	HSA241023	HSA241023	Homo sapi
c 112	41	3.9	10151	9	PTR270063	PTR270063	Pan trogl	188	39	3.7	10163	9	HSA241024	HSA241024	Homo sapi
c 113	41	3.9	16459	9	HSA297560	HSA297560	Homo sapi	189	39	3.7	10163	9	HSA241025	HSA241025	Homo sapi
c 114	41	3.9	90476	2	AF209070	AF209070	Homo sapi	190	39	3.7	10163	9	HSA241026	HSA241026	Homo sapi
c 115	41	3.9	111857	2	HSJ636FL13	HSJ636FL13	Homo sapi	191	39	3.7	10163	9	HSA241027	HSA241027	Homo sapi
c 116	41	3.9	119290	2	AC068449	AC068449	Homo sapi	192	39	3.7	10163	9	HSA241028	HSA241028	Homo sapi
c 117	41	3.9	121417	2	AL359374	AL359374	Homo sapi	193	39	3.7	10163	9	HSA241029	HSA241029	Homo sapi
c 118	41	3.9	121826	9	AL354935	AL354935	Human DNA	194	39	3.7	10163	9	HSA241030	HSA241030	Homo sapi
c 119	41	3.9	124874	9	HS888M10	HS888M10	Human DNA	195	39	3.7	10163	9	HSA241031	HSA241031	Homo sapi
c 120	41	3.9	130351	2	AC024911	AC024911	Homo sapi	196	39	3.7	10163	9	HSA241032	HSA241032	Homo sapi
c 121	41	3.9	130855	9	AC004089	AC004089	Homo sapi	197	39	3.7	10163	9	HSA241033	HSA241033	Homo sapi
c 122	41	3.9	133814	2	AC027550	AC027550	Homo sapi	198	39	3.7	10163	9	HSA241034	HSA241034	Homo sapi
c 123	41	3.9	140974	9	AC068533	AC068533	Homo sapi	199	39	3.7	10163	9	HSA241035	HSA241035	Homo sapi
c 124	41	3.9	143489	2	AL161637	AL161637	Homo sapi	200	39	3.7	10163	9	HSA241036	HSA241036	Homo sapi
125	41	3.9	144067	2	AL138798	AL138798	Homo sapi	201	39	3.7	10163	9	HSA241037	HSA241037	Homo sapi
126	41	3.9	150747	2	AC016230	AC016230	Homo sapi	202	39	3.7	10163	9	HSA241038	HSA241038	Homo sapi
127	41	3.9	150781	2	AC091961	AC091961	Homo sapi	203	39	3.7	10163	9	HSA241039	HSA241039	Homo sapi
c 128	41	3.9	150849	2	AC020890	AC020890	Homo sapi	204	39	3.7	10163	9	HSA241040	HSA241040	Homo sapi
c 129	41	3.9	150849	2	AC020890	AC020890	Homo sapi	205	39	3.7	10163	9	HSA241041	HSA241041	Homo sapi
c 130	41	3.9	155215	2	AL391220	AL391220	Homo sapi	206	39	3.7	10163	9	HSA241042	HSA241042	Homo sapi
c 131	41	3.9	156604	9	AL356865	AL356865	Human DNA	207	39	3.7	10163	9	HSA241043	HSA241043	Homo sapi
132	41	3.9	161109	2	AC021028	AC021028	Homo sapi	208	39	3.7	10163	9	HSA241044	HSA241044	Homo sapi
133	41	3.9	161609	2	AC011823	AC011823	Homo sapi	209	39	3.7	10163	9	HSA241045	HSA241045	Homo sapi
134	41	3.9	161999	2	AL133352	AL133352	Homo sapi	210	39	3.7	10163	9	HSA241046	HSA241046	Homo sapi
c 135	41	3.9	162539	2	AL390720	AL390720	Homo sapi	211	39	3.7	10163	9	HSA241047	HSA241047	Homo sapi
c 136	41	3.9	162609	9	AC021171	AC021171	Homo sapi	212	39	3.7	10163	9	HSA241048	HSA241048	Homo sapi
c 137	41	3.9	164264	2	AC034197	AC034197	Homo sapi	213	39	3.7	10163	9	HSA241049	HSA241049	Homo sapi
c 138	41	3.9	165484	2	AL590671	AL590671	Homo sapi	214	39	3.7	10163	9	HSA241050	HSA241050	Homo sapi
c 139	41	3.9	166430	2	AL157895	AL157895	Homo sapi	215	39	3.7	10163	9	HSA241051	HSA241051	Homo sapi
c 140	41	3.9	168865	2	AC093123	AC093123	Papio cyn	216	39	3.7	10163	9	HS241052	HS241052	Homo sapi
c 141	41	3.9	168865	2	AC093123	AC093123	Papio cyn								
c 142	41	3.9	168908	2	CNS07BEQ	AL445883	Homo sapi								
c 143	41	3.9	170352	2	AL358974	AL358974	Homo sapi								
	41	3.9	171530	9	AC083870	AC083870	Homo sapi								
	41	3.9	172170	2	AC018357	AC018357	Homo sapi								

217	39	3.7	10163	9	HSA241053	AJ241053	Homo sapi	290	38	3.7	198486	2	AC009118	AC009118	Homo sapi
218	39	3.7	10163	9	HSA241054	AJ241054	Homo sapi	291	38	3.7	227968	9	AF053356	AF053356	Homo sapi
219	39	3.7	10163	9	HSA241055	AJ241055	Homo sapi	292	37	3.6	10531	9	AC020902	AC020902	Homo sapi
220	39	3.7	10163	9	HSA241056	AJ241056	Homo sapi	c 293	37	3.6	15239	2	AC090678	AC090678	Homo sapi
221	39	3.7	10163	9	HSA241057	AJ241057	Homo sapi	c 294	37	3.6	24232	9	AL590383	AL590383	Human DNA
222	39	3.7	10163	9	HSA241058	AJ241058	Homo sapi	c 295	37	3.6	33458	9	HSDJ60101	AL109656	Human DNA
223	39	3.7	10163	9	HSA241059	AJ241059	Homo sapi	296	37	3.6	39322	9	AP001741	AP001741	Homo sapi
224	39	3.7	10163	9	HSA241060	AJ241060	Homo sapi	c 297	37	3.6	54666	9	AC073487	AC073487	Homo sapi
225	39	3.7	10163	9	HSA241061	AJ241061	Homo sapi	c 298	37	3.6	57000	9	AC069278	AC069278	Homo sapi
226	39	3.7	10163	9	HSA241062	AJ241062	Homo sapi	c 299	37	3.6	68688	2	AC027794	AC027794	Homo sapi
227	39	3.7	10163	9	HSA241063	AJ241063	Homo sapi	c 300	37	3.6	76165	2	AP000583	AP000583	Homo sapi
228	39	3.7	10163	9	HSA241064	AJ241064	Homo sapi	c 301	37	3.6	89626	9	AC004776	AC004776	Homo sapi
229	39	3.7	10163	9	HSA241065	AJ241065	Homo sapi	c 302	37	3.6	97835	9	HS292E10	Z93930	Human DNA s
230	39	3.7	10163	9	HSA241066	AJ241066	Homo sapi	303	37	3.6	100000	9	AP000132	AP000132	Homo sapi
231	39	3.7	10163	9	HSA241067	AJ241067	Homo sapi	304	37	3.6	100000	9	AP000210	AP000210	Homo sapi
232	39	3.7	10163	9	HSA241068	AJ241068	Homo sapi	c 305	37	3.6	105412	2	AC011458	AC011458	Homo sapi
233	39	3.7	10163	9	HSA241069	AJ241069	Homo sapi	306	37	3.6	114992	9	AC008772	AC008772	Homo sapi
234	39	3.7	10163	9	HSA241070	AJ241070	Homo sapi	c 307	37	3.6	117564	2	HSDJ160H5	AL096873	Homo sapi
235	39	3.7	10163	9	HSA241071	AJ241071	Homo sapi	c 308	37	3.6	120428	2	AP000721	AP000721	Homo sapi
236	39	3.7	10163	9	HSA241072	AJ241072	Homo sapi	c 309	37	3.6	120578	9	AC005887	AC005887	citb_173_
237	39	3.7	10163	9	HSA241073	AJ241073	Homo sapi	c 310	37	3.6	120608	9	HS469K11	AL449209	Homo sapi
238	39	3.7	10163	9	HSA241074	AJ241074	Homo sapi	c 311	37	3.6	127661	2	AP001261	AP001261	Homo sapi
239	39	3.7	10163	9	HSA241075	AJ241075	Homo sapi	312	37	3.6	130282	2	AC078973	AC078973	Homo sapi
240	39	3.7	10163	9	HSA241076	AJ241076	Homo sapi	313	37	3.6	132110	9	AP001610	AP001610	Homo sapi
241	39	3.7	10163	9	HSA241077	AJ241077	Homo sapi	c 314	37	3.6	133893	9	AC005005	AC005005	Homo sapi
242	39	3.7	10163	9	HSA241078	AJ241078	Homo sapi	c 315	37	3.6	134413	2	AC092932	AC092932	Homo sapi
243	39	3.7	10163	9	HSA241079	AJ241079	Homo sapi	c 316	37	3.6	142046	2	AC084024	AC084024	Homo sapi
244	39	3.7	10163	9	HSA241080	AJ241080	Homo sapi	c 317	37	3.6	142046	2	AC010630	AC010630	Homo sapi
245	39	3.7	10163	9	HSA241081	AJ241081	Homo sapi	318	37	3.6	143029	2	AC016319	AC016319	Homo sapi
246	39	3.7	10163	9	HSA241082	AJ241082	Homo sapi	c 319	37	3.6	143226	9	AL353802	AL353802	Human DNA
247	39	3.7	10163	9	HSA241083	AJ241083	Homo sapi	c 320	37	3.6	146017	2	AC027473	AC027473	Homo sapi
248	39	3.7	10163	9	HSA241084	AJ241084	Homo sapi	321	37	3.6	146847	9	AP000248	AP000248	Homo sapi
249	39	3.7	10163	9	HSA241085	AJ241085	Homo sapi	c 322	37	3.6	152448	2	AC008688	AC008688	Homo sapi
250	39	3.7	10163	9	HSA241086	AJ241086	Homo sapi	323	37	3.6	150221	9	AC004593	AC004593	Homo sapi
251	39	3.7	10163	9	HSA241087	AJ241087	Homo sapi	c 324	37	3.6	151734	2	AC044871	AC044871	Homo sapi
252	39	3.7	10163	9	HSA241088	AJ241088	Homo sapi	325	37	3.6	150893	2	AC021225	AC021225	Homo sapi
253	39	3.7	10163	9	HSA241089	AJ241089	Homo sapi	326	37	3.6	152448	2	AC021391	AC021391	Homo sapi
254	39	3.7	10163	9	HSA241090	AJ241090	Homo sapi	c 327	37	3.6	152607	2	AP003780	AP003780	Homo sapi
255	39	3.7	10163	9	HSA241091	AJ241091	Homo sapi	c 328	37	3.6	152679	2	AC024109	AC024109	Homo sapi
256	39	3.7	10164	9	HSA312532	AJ312532	Homo sapi	c 329	37	3.6	153297	2	AC041040	AC041040	Homo sapi
c 257	39	3.7	15143	9	HSAJ4862	AJ004862	Homo sapi	c 330	37	3.6	153788	2	AC023971	AC023971	Homo sapi
c 258	39	3.7	22773	9	AF107890	AF107890	Homo sapi	c 331	37	3.6	155488	2	AC008373	AC008373	Homo sapi
c 259	39	3.7	128011	9	HS333E23	Z82200	Human DNA s	c 332	37	3.6	155585	2	AC055727	AC055727	Homo sapi
260	39	3.7	163559	9	AC018635	AC018635	Homo sapi	c 333	37	3.6	156741	2	AL355867	AL355867	Homo sapi
c 261	39	3.7	166824	2	AC010655	AC010655	Homo sapi	334	37	3.6	159744	2	AC046139	AC046139	Homo sapi
262	39	3.7	173877	9	AC089982	AC089982	Homo sapi	c 335	37	3.6	160903	2	AP002776	AP002776	Homo sapi
263	39	3.7	174424	2	HSU52112	U52112	Homo sapien	c 336	37	3.6	161822	2	AC026306	AC026306	Homo sapi
264	39	3.7	174511	2	AC061979	AP003559	Homo sapi	337	37	3.6	162912	2	AC092954	AC092954	Homo sapi
265	39	3.7	174877	2	AP003559	AF069231	Homo sapi	c 338	37	3.6	163873	2	AC026058	AC026058	Homo sapi
266	39	3.7	177240	2	AC069231	AF129075	Homo sapi	c 339	37	3.6	163873	2	AC026080	AC026080	Homo sapi
267	39	3.7	189413	9	AL354735	AL354735	Human DNA	c 340	37	3.6	166014	2	AC010223	AC010223	Homo sapi
c 268	39	3.7	189821	2	AP003108	AL354735	Human DNA	c 341	37	3.6	167368	9	AC010223	AC010223	Homo sapi
269	39	3.7	204267	2	AC019071	AP003108	Homo sapi	c 342	37	3.6	167728	2	AC055870	AC055870	Homo sapi
c 270	39	3.7	214526	2	AP001809	AC020916	Homo sapi	c 343	37	3.6	167914	2	AC055870	AC055870	Homo sapi
c 271	39	3.7	215881	2	AC001809	AP001809	Homo sapi	c 344	37	3.6	171386	2	AP001198	AP001198	Homo sapi
272	39	3.7	340000	9	HS21C049	AL163249	Homo sapi	c 345	37	3.6	171566	2	AP001885	AP001885	Homo sapi
273	39	3.7	66788	9	AL353653	AL353653	Human DNA	c 346	37	3.6	172476	9	AC007736	AC007736	Homo sapi
c 274	38	3.7	84661	9	AL353695	AL353695	Human DNA	c 347	37	3.6	172570	9	AC006064	AC006064	Homo sapi
c 275	38	3.7	141378	2	AC069281	AC069281	Homo sapi	c 348	37	3.6	175126	2	AL590094	AL590094	Homo sapi
276	38	3.7	153289	2	AC013533	AC013533	Homo sapi	c 349	37	3.6	175461	2	AP001201	AP001201	Homo sapi
c 277	38	3.7	153289	2	AC011156	AC011156	Homo sapi	350	37	3.6	179259	2	AL157392	AL157392	Homo sapi
c 278	38	3.7	156573	2	AC073075	AC073075	Homo sapi	c 351	37	3.6	179507	2	AC015976	AC015976	Homo sapi
c 279	38	3.7	159387	2	AC010162	AC010162	Homo sapi	c 352	37	3.6	181671	9	AC007040	AC007040	Homo sapi
c 280	38	3.7	163781	9	AC058784	AC058784	Homo sapi	c 353	37	3.6	184662	9	AC025436	AC025436	Homo sapi
c 281	38	3.7	164306	2	AP002785	AP002785	Homo sapi	c 354	37	3.6	184978	2	AL590374	AL590374	Homo sapi
c 282	38	3.7	164306	2	AP002785	AL591647	Homo sapi	355	37	3.6	185134	9	AC007878	AC007878	Homo sapi
c 283	38	3.7	164388	2	AL591647	AL591647	Homo sapi	356	37	3.6	188165	9	CNS00003	AL049781	Human chr
284	38	3.7	168156	2	AL590648	AL590648	Homo sapi	c 357	37	3.6	189317	9	HSJ190J20	AL050335	Human DNA
285	38	3.7	173829	2	AL451077	AL451077	Homo sapi	358	37	3.6	190682	2	AC024293	AC024293	Homo sapi
286	38	3.7	186418	2	AC026898	AC026898	Homo sapi	c 359	37	3.6	191853	2	AC084151	AC084151	Homo sapi
287	38	3.7	190872	2	AC015799	AC015799	Homo sapi	c 360	37	3.6	192453	2	AC023920	AC023920	Homo sapi
c 288	38	3.7	193026	2	AC010158	AC010158	Homo sapi	c 361	37	3.6	193212	9	AC007405	AC007405	Homo sapi
c 289	38	3.7	194681	2	AC084057	AC084057	Homo sapi	c 362	37	3.6	196556	2	AC018382	AC018382	Homo sapi

c 363	37	3.6	200379	2	AC023055	AC023055 Homo sapi	c 436	36	3.5	219820	9	AC008738	Homo sapi
c 364	37	3.6	201395	2	AC040165	AC040165 Homo sapi	c 437	36	3.5	230000	9	AF243527	Homo sapi
c 365	37	3.6	207653	2	AL355614	AL355614 Homo sapi	438	35	3.4	10141	9	PPY270095	Pongo pyg
c 366	37	3.6	214989	2	AC019115	AC019115 Homo sapi	439	35	3.4	10142	9	PPY312509	Pongo pyg
c 367	37	3.6	216300	2	AC069335	AC069335 Homo sapi	440	35	3.4	10143	9	PPY312511	Pongo pyg
c 368	37	3.6	221420	2	AC011822	AC011822 Homo sapi	441	35	3.4	10143	9	PPY312515	Pongo pyg
c 369	37	3.6	234844	2	AL359836	AL359836 Homo sapi	442	35	3.4	10144	9	PPY312513	Pongo pyg
c 370	37	3.6	269711	2	AP000408	AP000408 Homo sapi	443	35	3.4	10144	9	PPY312520	Pongo pyg
c 371	37	3.6	275159	9	U82670	U82670 Homo sapien	c 444	35	3.4	36339	9	HSB33B7	Human DNA s
c 372	37	3.6	340000	9	AP001711	AP001711 Homo sapi	445	35	3.4	47198	2	AC090252	Homo sapi
c 373	36	3.5	6500	9	AF095743	AF095743 Homo sapi	446	35	3.4	73360	9	HSU66083	Human conti
c 374	36	3.5	7680	9	AF251125	AF251125 Homo sapi	447	35	3.4	74853	9	AC007900	Homo sapi
c 375	36	3.5	10142	9	PPY312508	PPY312508 Pongo pyg	c 448	35	3.4	76241	9	HSDJ336M4	Human DNA
c 376	36	3.5	10142	9	PPY312512	AJ312512 Pongo pyg	c 449	35	3.4	85116	9	HS324M8	Human DNA
c 377	36	3.5	10142	9	PPY312516	AJ312516 Pongo pyg	c 450	35	3.4	92510	9	HS390B3	Human DNA s
c 378	36	3.5	10142	9	PPY312517	AJ312517 Pongo pyg	c 451	35	3.4	103863	9	AC006596	Homo sapi
c 379	36	3.5	10142	9	PPY312518	AJ312518 Pongo pyg	c 452	35	3.4	111183	9	AC008950	Homo sapi
c 380	36	3.5	10142	9	PPY312519	AJ312519 Pongo pyg	453	35	3.4	112314	9	AC004002	Human BAC
c 381	36	3.5	10144	9	PPY312510	AJ312510 Pongo pyg	454	35	3.4	113297	9	AL138762	Human DNA
c 382	36	3.5	10144	9	PPY312514	AJ312514 Pongo pyg	455	35	3.4	114026	9	AL1365201	Human DNA
c 383	36	3.5	39540	9	AC000079	AC000079 Homo sapi	456	35	3.4	122839	9	AL161898	Human DNA
c 384	36	3.5	43934	9	AC000068	AC000068 Homo sapi	c 457	35	3.4	128880	2	AL136458	Homo sapi
c 385	36	3.5	46581	9	AC008982	AC008982 Homo sapi	458	35	3.4	129098	9	HS46H23	Human DNA s
c 386	36	3.5	71202	2	AL603649	AL603649 Homo sapi	459	35	3.4	129158	2	AC090170	Homo sapi
c 387	36	3.5	73239	9	HS373H7	Z99774 Human DNA s	460	35	3.4	135734	2	AC002346	Homo sapi
c 388	36	3.5	93196	2	AC022085	AC022085 Homo sapi	c 461	35	3.4	138538	9	AC011445	Homo sapi
c 389	36	3.5	98697	9	AC004854	AC004854 Homo sapi	c 462	35	3.4	139015	2	AC068016	Homo sapi
c 390	36	3.5	99336	2	AC007411	AC007411 Homo sapi	c 463	35	3.4	145947	9	AL353588	Human DNA
c 391	36	3.5	105866	2	AC021602	AC021602 Homo sapi	c 464	35	3.4	147708	9	HS179M20	Human DNA s
c 392	36	3.5	109124	2	AC011483	AC011483 Homo sapi	c 465	35	3.4	148271	9	AL392046	Human DNA
c 393	36	3.5	113872	9	HS57G9	Z95116 Human DNA s	c 466	35	3.4	150435	2	AC069579	Homo sapi
c 394	36	3.5	128266	9	AC008623	AC008623 Homo sapi	c 467	35	3.4	155186	2	AC068373	Homo sapi
c 395	36	3.5	134929	9	AC004655	AC004655 Homo sapi	468	35	3.4	157095	2	AL355994	Homo sapi
c 396	36	3.5	137701	2	AC079088	AC079088 Homo sapi	469	35	3.4	157365	9	AC025754	Homo sapi
c 397	36	3.5	142334	2	AC073185	AC073185 Homo sapi	c 470	35	3.4	160725	2	HS919B11	Homo sapi
c 398	36	3.5	145447	2	AC012651	AC012651 Homo sapi	c 471	35	3.4	162486	2	AC044779	Homo sapi
c 399	36	3.5	148540	9	HS212P9	AL009181 Human DNA	472	35	3.4	164034	9	AC011489	Homo sapi
c 400	36	3.5	149765	9	AC004984	AC004984 Homo sapi	c 473	35	3.4	169652	2	AC090668	Homo sapi
c 401	36	3.5	152395	2	AC079168	AC079168 Homo sapi	474	35	3.4	170790	2	AC025609	Homo sapi
c 402	36	3.5	153640	9	AL445669	AL445669 Human DNA	c 475	35	3.4	171206	2	AC026573	Homo sapi
c 403	36	3.5	153875	9	AC003682	AC003682 Homo sapi	476	35	3.4	172090	2	AC018477	Homo sapi
c 404	36	3.5	154848	9	AC003992	AC002992 Homo sapi	477	35	3.4	178574	2	AL591026	Homo sapi
c 405	36	3.5	155663	2	AC069238	AC069238 Homo sapi	c 478	35	3.4	178953	2	AC026626	Homo sapi
c 406	36	3.5	156580	9	AP003493	AP003493 Homo sapi	c 479	35	3.4	179667	2	AC025681	Homo sapi
c 407	36	3.5	157579	9	AC018639	AC018639 Human Chr	c 480	35	3.4	180615	2	AC053543	Homo sapi
c 408	36	3.5	159693	2	AC016727	AC016727 Homo sapi	c 481	35	3.4	186261	9	AC016602	Homo sapi
c 409	36	3.5	159854	2	AC069524	AC069524 Homo sapi	c 482	35	3.4	186686	2	AL162739	Homo sapi
c 410	36	3.5	159873	2	AC024023	AC024023 Homo sapi	483	35	3.4	189010	2	AC092835	Homo sapi
c 411	36	3.5	163320	2	AC027294	AC027294 Homo sapi	c 484	35	3.4	193323	2	AC073842	Homo sapi
c 412	36	3.5	167335	2	AC023527	AC023527 Homo sapi	c 485	35	3.4	193514	9	AC010530	Homo sapi
c 413	36	3.5	168843	33	AL356972	AL356972 Human DNA	486	35	3.4	197326	2	AC022966	Homo sapi
c 414	36	3.5	169997	9	AC008440	AC008440 Homo sapi	c 487	35	3.4	199019	2	AL359764	Homo sapi
c 415	36	3.5	172462	2	AL355529	AL355529 Homo sapi	488	35	3.4	201883	2	AC009972	Homo sapi
c 416	36	3.5	173396	2	AL390236	AL390236 Homo sapi	489	35	3.4	203650	2	AC092296	Homo sapi
c 417	36	3.5	175223	9	AC004617	AC004617 Homo sapi	c 490	35	3.4	208337	2	AC068400	Homo sapi
c 418	36	3.5	175585	2	AC023590	AC023590 Homo sapi	491	35	3.4	210672	2	HS179115	Human sapien
c 419	36	3.5	178651	9	CNS057DY	AL358275 Human chr	c 492	34	3.3	26459	9	AL390854	Human DNA
c 420	36	3.5	180195	2	AC027089	AC027089 Homo sapi	493	34	3.3	28505	9	HSBA775A3	Human DNA
c 421	36	3.5	182776	9	AC016748	AC016748 Homo sapi	494	34	3.3	48730	9	AL136301	Human DNA
c 422	36	3.5	183600	2	AC093131	AC093131 Papio cyn	495	34	3.3	69213	9	AC010191	Homo sapi
c 423	36	3.5	184414	2	AC092135	AC092135 Homo sapi	496	34	3.3	82412	2	AC040940	Homo sapi
c 424	36	3.5	188934	2	AC072040	AC072040 Homo sapi	497	34	3.3	94742	9	AC006263	Homo sapi
c 425	36	3.5	191422	2	AC022318	AC022318 Homo sapi	c 498	34	3.3	99593	9	HSJ857M17	Human DNA
c 426	36	3.5	194361	2	AC092017	AC092017 Homo sapi	499	34	3.3	110000	2	AL355364_1	Continuation (2 of
c 427	36	3.5	195969	2	AC017107	AC017107 Homo sapi	500	34	3.3	140092	9	AL162615	Human DNA
c 428	36	3.5	196844	2	AC067730	AC067730 Homo sapi	501	34	3.3	145966	9	AC006451	Homo sapi
c 429	36	3.5	199014	2	AC073849	AC073849 Homo sapi	c 502	34	3.3	146028	2	AC021573	Homo sapi
c 430	36	3.5	199019	2	AL359764	AL359764 Homo sapi	c 503	34	3.3	152283	9	AC024083	Homo sapi
c 431	36	3.5	201123	2	AL139410	AL139410 Homo sapi	c 504	34	3.3	156005	2	AL158156	Homo sapi
c 432	36	3.5	203200	9	AC008744	AC008744 Homo sapi	505	34	3.3	157633	9	AC008687	Homo sapi
c 433	36	3.5	216935	2	AC083813	AC083813 Homo sapi	506	34	3.3	158297	2	AL157820	Homo sapi
c 434	36	3.5	217346	2	AC027602	AC027602 Homo sapi	c 507	34	3.3	158594	2	AC087768	Homo sapi
c 435	36	3.5	217615	9	AC011005	AC011005 Homo sapi	508	34	3.3	159972	33	AC022932	Homo sapi

c 509	34	3.3	159983	2	AC013713	AC013713 Homo sapi
c 510	34	3.3	162847	2	AC025980	AC025980 Homo sapi
c 511	34	3.3	164917	2	AL592064	AL592064 Homo sapi
512	34	3.3	165611	2	AC011694	AC011694 Homo sapi
513	34	3.3	166354	2	AC078984	AC078984 Homo sapi
c 514	34	3.3	167563	9	AC069280	AC069280 Homo sapi
c 515	34	3.3	168247	9	AC005014	AC005014 Homo sapi
c 516	34	3.3	171007	2	AC090421	AC090421 Homo sapi
517	34	3.3	171176	2	AC060796	AC060796 Homo sapi
518	34	3.3	171338	2	AL157875	AL157875 Homo sapi
c 519	34	3.3	171998	2	AC092648	AC092648 Homo sapi
520	34	3.3	174761	2	AL391650	AL391650 Homo sapi
c 521	34	3.3	178630	2	AC016219	AC016219 Homo sapi
c 522	34	3.3	179085	9	AC011611	AC011611 Homo sapi
c 523	34	3.3	179260	2	AC084291	AC084291 Homo sapi
524	34	3.3	181891	9	AC015987	AC015987 Homo sapi
c 525	34	3.3	181914	9	AC009154	AC009154 Homo sapi
526	34	3.3	182540	2	AC016916	AC016916 Homo sapi
c 527	34	3.3	183451	9	AC005972	AC005972 Homo sapi
528	34	3.3	185380	2	AC016863	AC016863 Homo sapi
c 529	34	3.3	189938	2	AC023601	AC023601 Homo sapi
c 530	34	3.3	192254	2	AC010741	AC010741 Homo sapi
531	34	3.3	192760	2	AC083928	AC083928 Homo sapi
c 532	34	3.3	195439	2	AL355586	AL355586 Homo sapi
c 533	34	3.3	197055	2	AL390774	AL390774 Homo sapi
534	34	3.3	199752	2	AC018475	AC018475 Homo sapi
535	34	3.3	210096	2	AL589942	AL589942 Homo sapi
c 536	34	3.3	214653	2	AL161790	AL161790 Homo sapi
c 537	34	3.3	215584	2	AL390779	AL390779 Homo sapi
538	34	3.3	231260	2	AL160172	AL160172 Homo sapi
c 540	34	3.3	271144	9	HSXDPB	HSXDPB Homo sapi
541	33	3.2	752	9	AF205538	AF205538 Homo sapi
c 542	33	3.2	5507	9	BC008313	BC008313 Homo sapi
c 543	33	3.2	32198	9	AF229163	AF229163 Homo sapi
c 544	33	3.2	36800	9	HS313D11	HS313D11 Homo sapi
545	33	3.2	49386	2	AC021257	AC021257 Homo sapi
c 546	33	3.2	73029	9	HSBA329J7	HSBA329J7 Homo sapi
c 547	33	3.2	76200	9	AC008616	AC008616 Homo sapi
c 548	33	3.2	80419	9	AL139826	AL139826 Homo sapi
549	33	3.2	83798	9	HS1013A22	HS1013A22 Homo sapi
550	33	3.2	86453	9	HS1049G16	HS1049G16 Homo sapi
c 551	33	3.2	87375	9	AP0000289	AP0000289 Homo sapi
c 552	33	3.2	89016	2	AP003326	AP003326 Homo sapi
c 553	33	3.2	90248	2	AF038458	AF038458 Homo sapi
c 554	33	3.2	92797	9	AC000066	AC000066 Homo sapi
555	33	3.2	94960	9	AP000042	AP000042 Homo sapi
c 556	33	3.2	100000	9	AP000049	AP000049 Homo sapi
c 557	33	3.2	100000	9	AP0000110	AP0000110 Homo sapi
c 558	33	3.2	100000	9	AP000116	AP000116 Homo sapi
c 559	33	3.2	100000	9	AP000186	AP000186 Homo sapi
c 560	33	3.2	100000	9	AP000192	AP000192 Homo sapi
c 561	33	3.2	100000	9	AP000503	AP000503 Homo sapi
c 562	33	3.2	102151	9	HS979N1	HS979N1 Homo sapi
563	33	3.2	108523	9	AC000159	AC000159 Homo sapi
c 564	33	3.2	109867	9	AC004898	AC004898 Homo sapi
565	33	3.2	110633	9	AL353710	AL353710 Homo sapi
566	33	3.2	111035	9	AC022405	AC022405 Homo sapi
567	33	3.2	111990	9	AL360294	AL360294 Homo sapi
c 568	33	3.2	112309	2	AP000667	AP000667 Homo sapi
c 569	33	3.2	118030	2	AP000667	AP000667 Homo sapi
c 570	33	3.2	121442	2	AL137860	AL137860 Homo sapi
571	33	3.2	121442	2	AL137860	AL137860 Homo sapi
c 572	33	3.2	123530	2	AL356369	AL356369 Homo sapi
c 573	33	3.2	125242	9	AP000311	AP000311 Homo sapi
c 574	33	3.2	125291	9	AP000646	AP000646 Homo sapi
575	33	3.2	125350	2	AC020768	AC020768 Homo sapi
576	33	3.2	125350	2	AC020768	AC020768 Homo sapi
577	33	3.2	128230	9	HS313L4	HS313L4 Homo sapi
578	33	3.2	128559	9	AC004677	AC004677 Homo sapi
579	33	3.2	128899	9	AL391139	AL391139 Homo sapi
580	33	3.2	129502	9	HSDJ726C3	HSDJ726C3 Homo sapi
c 581	33	3.2	130467	9	HS798A17	HS798A17 Homo sapi
c 582	33	3.2	131314	2	AP000817	AP000817 Homo sapi
583	33	3.2	133683	2	AC084857	AC084857 Homo sapi
584	33	3.2	134515	9	HS354J5	HS354J5 Homo sapi
585	33	3.2	134601	2	AL450106	AL450106 Homo sapi
c 586	33	3.2	136572	9	AL133384	AL133384 Human DNA
c 587	33	3.2	137833	9	CNS01DV6	CNS01DV6 Human chr
588	33	3.2	138094	2	AC009417	AC009417 Homo sapi
c 589	33	3.2	138627	2	AP000591	AP000591 Homo sapi
c 590	33	3.2	140606	2	AL353761	AL353761 Homo sapi
c 591	33	3.2	141989	2	AC069439	AC069439 Homo sapi
592	33	3.2	142092	9	AF139813	AF139813 Homo sapi
c 593	33	3.2	143769	9	AC004914	AC004914 Homo sapi
c 594	33	3.2	143813	9	AC002994	AC002994 Homo sapi
595	33	3.2	144200	2	AC022638	AC022638 Homo sapi
c 596	33	3.2	144676	9	HS390013	HS390013 Human DNA s
c 597	33	3.2	145414	9	HSA392M18	HSA392M18 Human DNA
598	33	3.2	147094	2	AC012027	AC012027 Homo sapi
599	33	3.2	149425	2	AC013320	AC013320 Homo sapi
c 600	33	3.2	149733	2	AC087568	AC087568 Pan trogl
601	33	3.2	149765	9	AC004984	AC004984 Homo sapi
602	33	3.2	153130	2	AC091980	AC091980 Homo sapi
603	33	3.2	153284	2	AL136233	AL136233 Homo sapi
c 604	33	3.2	156657	2	AL139260	AL139260 Homo sapi
605	33	3.2	156795	2	AC023537	AC023537 Homo sapi
606	33	3.2	157662	9	AL139035	AL139035 Human DNA
607	33	3.2	158060	2	AC024423	AC024423 Homo sapi
c 608	33	3.2	158287	2	AP002961	AP002961 Homo sapi
c 609	33	3.2	158475	2	CNS01DXB	CNS01DXB Homo sapi
610	33	3.2	158802	2	AC015589	AC015589 Homo sapi
c 611	33	3.2	158802	2	AC015589	AC015589 Homo sapi
c 612	33	3.2	158885	33	AC007847	AC007847 Homo sapi
613	33	3.2	159624	2	AC011021	AC011021 Homo sapi
614	33	3.2	159635	2	AC011009	AC011009 Homo sapi
615	33	3.2	160156	2	AL591852	AL591852 Homo sapi
c 616	33	3.2	161031	9	AC079610	AC079610 Homo sapi
617	33	3.2	161548	9	AP001929	AP001929 Homo sapi
c 618	33	3.2	162453	2	AC087778	AC087778 Pan trogl
619	33	3.2	162589	2	AC015717	AC015717 Homo sapi
620	33	3.2	162984	2	AL590642	AL590642 Homo sapi
621	33	3.2	163682	9	HSBJ71117	HSBJ71117 Human DNA
c 622	33	3.2	164426	2	AC025154	AC025154 Homo sapi
c 623	33	3.2	166118	9	AL355495	AL355495 Human DNA
c 624	33	3.2	166774	9	AC005225	AC005225 Homo sapi
625	33	3.2	166894	2	AL161787	AL161787 Homo sapi
626	33	3.2	168119	2	AC068686	AC068686 Homo sapi
c 627	33	3.2	168210	9	AC018719	AC018719 Homo sapi
c 628	33	3.2	168544	9	AL133344	AL133344 Human DNA
629	33	3.2	169223	2	AL353675	AL353675 Homo sapi
c 630	33	3.2	169812	9	AL391315	AL391315 Human DNA
c 631	33	3.2	169863	2	AL591467	AL591467 Homo sapi
c 632	33	3.2	169984	2	AL161789	AL161789 Homo sapi
633	33	3.2	170008	9	HSA010770	HSA010770 Homo sapi
c 634	33	3.2	170281	2	AL592290	AL592290 Homo sapi
c 635	33	3.2	170591	9	AC026410	AC026410 Homo sapi
c 636	33	3.2	171034	2	AP001325	AP001325 Homo sapi
637	33	3.2	171759	2	AC073650	AC073650 Homo sapi
c 638	33	3.2	171980	2	AP000928	AP000928 Homo sapi
639	33	3.2	172105	2	AC021197	AC021197 Homo sapi
c 640	33	3.2	172138	2	AC064796	AC064796 Homo sapi
c 641	33	3.2	173585	9	AC090527	AC090527 Homo sapi
c 642	33	3.2	174733	2	AC016386	AC016386 Homo sapi
643	33	3.2	175043	2	AC024246	AC024246 Homo sapi
644	33	3.2	175307	2	AC093246	AC093246 Homo sapi
645	33	3.2	176584	2	AC027812	AC027812 Homo sapi
646	33	3.2	177426	2	AC021222	AC021222 Homo sapi
c 647	33	3.2	178087	9	AC005089	AC005089 Homo sapi
648	33	3.2	178089	2	AC093133	AC093133 Papio cyn
c 649	33	3.2	178460	2	AL139040	AL139040 Homo sapi
650	33	3.2	179182	2	AC092965	AC092965 Homo sapi
651	33	3.2	179454	2	AC024177	AC024177 Homo sapi
c 652	33	3.2	180038	2	AC068393	AC068393 Homo sapi
c 653	33	3.2	180283	9	AF134726	AF134726 Homo sapi
c 654	33	3.2	180849	2	AC023996	AC023996 Homo sapi

c 655	33	3. 2	181054	2	AC025019	Homo sapi
c 656	33	3. 2	181670	2	AP003041	Homo sapi
657	33	3. 2	182295	2	AP001501	Homo sapi
658	33	3. 2	182768	2	AC093189	Pan trogl
659	33	3. 2	183099	2	AC034144	Homo sapi
c 660	33	3. 2	183624	2	AC019101	Homo sapi
661	33	3. 2	183700	2	AC068612	Homo sapi
662	33	3. 2	183991	9	AC007391	Homo sapi
c 663	33	3. 2	185002	2	AC015989	Homo sapi
c 664	33	3. 2	185859	2	AC068026	Homo sapi
665	33	3. 2	186635	2	AC051645	Homo sapi
666	33	3. 2	188481	9	AL138836	Human DNA
667	33	3. 2	190663	2	AC068581	Homo sapi
668	33	3. 2	191450	2	AC093138	Pan trogl
c 669	33	3. 2	191754	9	AC021016	Homo sapi
670	33	3. 2	194431	2	AC093153	Homo sapi
c 671	33	3. 2	194658	2	AC020911	Homo sapi
672	33	3. 2	195761	33	AL357555	Human DNA
c 673	33	3. 2	196080	9	AC004228	Homo sapi
674	33	3. 2	196686	9	AC010328	Homo sapi
c 675	33	3. 2	197837	2	AP001397	Homo sapi
676	33	3. 2	198199	2	AL445673	Homo sapi
c 677	33	3. 2	199660	2	AC026469	Homo sapi
c 678	33	3. 2	200026	2	AC090368	Homo sapi
679	33	3. 2	202228	2	AL356390	Homo sapi
680	33	3. 2	202550	2	AL391832	Homo sapi
681	33	3. 2	203743	2	AC048337	Homo sapi
682	33	3. 2	204037	2	AC015881	Homo sapi
683	33	3. 2	204590	2	AC093130	Papio cyn
c 684	33	3. 2	204590	2	AC093130	Papio cyn
685	33	3. 2	206773	2	AL356218	Homo sapi
686	33	3. 2	207418	2	AC008676	Homo sapi
687	33	3. 2	210348	2	AC025845	Homo sapi
c 688	33	3. 2	213444	2	AC012202	Homo sapi
689	33	3. 2	221138	2	AC010454	Homo sapi
c 690	33	3. 2	237231	2	AC008382	Homo sapi
691	33	3. 2	240455	2	AC010290	Homo sapi
c 692	33	3. 2	256073	9	AE006464	Homo sapi
c 693	33	3. 2	268238	9	AC008372	Homo sapi
694	33	3. 2	300503	2	AC010260	Homo sapi
c 695	33	3. 2	340000	9	AP001716	Homo sapi
c 696	33	3. 2	340000	9	AP001718	Homo sapi
697	32	3. 1	67276	2	AC026622	Homo sapi
698	32	3. 1	78514	2	AL356496	Homo sapi
699	32	3. 1	78539	9	AP000313	Homo sapi
700	32	3. 1	83744	9	AP000194	Homo sapi
c 701	32	3. 1	96067	9	AP000953	Homo sapi
c 702	32	3. 1	100000	9	AB020866	Homo sapi
c 703	32	3. 1	102308	9	AC005231	Homo sapi
704	32	3. 1	107172	9	HS737M10	Human DNA
c 705	32	3. 1	108010	2	AL139258	Homo sapi
706	32	3. 1	111183	9	AC008950	Homo sapi
c 707	32	3. 1	114929	9	AP000050	Homo sapi
c 708	32	3. 1	118929	2	AC025976	Homo sapi
c 709	32	3. 1	119638	2	AC024575	Homo sapi
710	32	3. 1	130526	2	AP000715	Homo sapi
c 711	32	3. 1	141678	2	AL355816	Homo sapi
c 712	32	3. 1	143291	9	AL137792	Human DNA
713	32	3. 1	148312	2	AP001886	Homo sapi
714	32	3. 1	148355	2	AC027317	Homo sapi
715	32	3. 1	151516	9	AP000117	Homo sapi
c 716	32	3. 1	152464	2	AC027421	Homo sapi
c 717	32	3. 1	156897	2	AC018782	Homo sapi
c 718	32	3. 1	156929	2	AC026854	Homo sapi
719	32	3. 1	159044	2	AC025269	Homo sapi
c 720	32	3. 1	159769	2	AC007337	Homo sapi
c 721	32	3. 1	161155	9	AF017104	Homo sapi
c 722	32	3. 1	162593	2	AC022576	Homo sapi
c 723	32	3. 1	163229	9	AC009039	Homo sapi
724	32	3. 1	165866	2	AC087273	Homo sapi
725	32	3. 1	166297	2	AC011271	Homo sapi
726	32	3. 1	169357	2	AC010545	Homo sapi
727	32	3. 1	171077	2	AP001273	Homo sapi

c 728	32	3. 1	171250	2	AL513202	Homo sapi
729	32	3. 1	172581	9	AC009319	Homo sapi
c 730	32	3. 1	176495	2	AC087215	Papio cyn
c 731	32	3. 1	178229	2	AC010611	Homo sapi
732	32	3. 1	179525	9	AC087252	Papio cyn
733	32	3. 1	183338	2	AC083775	Homo sapi
734	32	3. 1	185475	2	AL159176	Homo sapi
c 735	32	3. 1	185838	2	AC036191	Homo sapi
736	32	3. 1	186261	9	AC016602	Homo sapi
c 737	32	3. 1	186396	2	AL359268	Homo sapi
c 738	32	3. 1	188324	2	AC087625	Homo sapi
739	32	3. 1	194173	9	CNS01RGW	Human chr
c 740	32	3. 1	198921	2	AC022250	Homo sapi
741	32	3. 1	208928	9	AC087729	Pan trogl
742	32	3. 1	210130	2	AC027110	Homo sapi
c 743	32	3. 1	210232	9	AC010899	Homo sapi
c 744	32	3. 1	217668	2	AC020561	Homo sapi
c 745	32	3. 1	237613	2	AC025769	Homo sapi
c 746	32	3. 1	340000	9	AP001684	Homo sapi
747	32	3. 1	340000	9	AP001718	Homo sapi
c 748	31	3. 0	38031	9	AC005764	Homo sapi
749	31	3. 0	58880	2	AC025855	Homo sapi
750	31	3. 0	59012	9	HSABLGR2	Human ABL g
c 751	31	3. 0	60009	2	AC023436	Homo sapi
c 752	31	3. 0	61957	9	AL135922	Human DNA
753	31	3. 0	66867	2	AC073170	Homo sapi
c 754	31	3. 0	66867	2	AC073170	Homo sapi
755	31	3. 0	68689	9	HSA278581	Homo sapi
c 756	31	3. 0	77405	2	AL353634	Homo sapi
757	31	3. 0	84544	2	AC009012	Homo sapi
758	31	3. 0	84710	9	HSBA358N2	Human DNA
759	31	3. 0	86914	2	AC011336	Homo sapi
c 760	31	3. 0	88217	2	AC006407	Homo sapi
761	31	3. 0	89264	9	AL353896	Human DNA
c 762	31	3. 0	102714	9	AP000692	Homo sapi
c 763	31	3. 0	110736	2	AL162741	Homo sapi
c 764	31	3. 0	118555	2	AL139287	Homo sapi
765	31	3. 0	121703	9	AC006452	Homo sapi
766	31	3. 0	129252	9	HSU91327	Human chrom
767	31	3. 0	130609	2	AC023217	Homo sapi
768	31	3. 0	133008	2	AC011375	Homo sapi
c 769	31	3. 0	134965	9	AL354923	Human DNA
c 770	31	3. 0	135946	2	AC008915	Homo sapi
c 771	31	3. 0	136131	9	AC000360	Homo sapi
c 772	31	3. 0	140377	2	AL355860	Homo sapi
773	31	3. 0	145140	2	AC073577	Homo sapi
774	31	3. 0	145190	2	AC008730	Homo sapi
c 775	31	3. 0	145642	2	AC027805	Homo sapi
776	31	3. 0	148285	2	AL451053	Homo sapi
c 777	31	3. 0	149411	2	AC019029	Homo sapi
778	31	3. 0	150159	2	AC018791	Homo sapi
779	31	3. 0	152878	2	AC079190	Homo sapi
c 780	31	3. 0	154028	9	AL136529	Human DNA
781	31	3. 0	154115	9	AC010878	Homo sapi
782	31	3. 0	154962	2	AL358153	Homo sapi
783	31	3. 0	155092	2	AC012073	Homo sapi
784	31	3. 0	157961	33	AC026051	Homo sapi
c 785	31	3. 0	157986	2	AC025860	Homo sapi
c 786	31	3. 0	158787	9	AC010886	Homo sapi
c 787	31	3. 0	158823	2	AC020759	Homo sapi
c 788	31	3. 0	158977	9	AP002469	Homo sapi
c 789	31	3. 0	159420	2	AC009017	Homo sapi
c 790	31	3. 0	159887	2	AC067823	Homo sapi
791	31	3. 0	160625	9	AC009311	Homo sapi
792	31	3. 0	161201	2	AC025673	Homo sapi
793	31	3. 0	161361	9	AC090939	Homo sapi
c 794	31	3. 0	161866	2	AC010614	Homo sapi
c 795	31	3. 0	161963	2	AL353622	Homo sapi
c 796	31	3. 0	161988	2	AC008728	Homo sapi
797	31	3. 0	162771	2	AL358173	Homo sapi
798	31	3. 0	163490	2	AC079354	Homo sapi
799	31	3. 0	164447	2	AC069148	Homo sapi
800	31	3. 0	164805	2	AC020978	Homo sapi

801	31	3.0	164945	9	AC089987	AC089987 Homo sapi	874	30	2.9	44183	2	AC093229	AC093229 Homo sapi
c 802	31	3.0	165110	9	AL450311	Human DNA	c 875	30	2.9	44183	2	AC093229	AC093229 Homo sapi
803	31	3.0	166234	9	HS273P12	Human DNA	876	30	2.9	47812	9	AC005351	AC005351 Homo sapi
804	31	3.0	167025	2	AC079866	Homo sapi	c 877	30	2.9	59847	9	AL136112	Human DNA
c 805	31	3.0	167880	9	HU091323	Human Chrom	c 878	30	2.9	63609	9	AL133343	Human DNA
c 806	31	3.0	169032	2	AC061965	Homo sapi	c 879	30	2.9	66741	9	HS460J8	Human DNA
c 807	31	3.0	169181	2	AC023303	Homo sapi	880	30	2.9	80600	9	AL121903	Human DNA
c 808	31	3.0	170000	2	AC004524	Homo sapi	881	30	2.9	92460	9	AL136228	Human DNA
c 809	31	3.0	170896	2	AC011010	Homo sapi	882	30	2.9	93165	9	AL357974	Human DNA
c 810	31	3.0	171757	9	CNS01DSO	Human chr	883	30	2.9	94673	2	AC009018	Homo sapi
c 811	31	3.0	172079	9	AL356095	Human DNA	c 884	30	2.9	98180	2	AC091867	Homo sapi
c 812	31	3.0	172376	2	AL353749	Homo sapi	c 885	30	2.9	98595	9	AP001329	Homo sapi
c 813	31	3.0	173109	2	AC091400	Pan trogl	c 886	30	2.9	101389	2	AL513365	Homo sapi
c 814	31	3.0	173229	2	AC091504	Pan trogl	c 887	30	2.9	101719	2	AL357495	Homo sapi
c 815	31	3.0	173341	2	AC021954	Homo sapi	c 888	30	2.9	102119	9	AL139005	Human DNA
816	31	3.0	173463	9	AL359092	Human DNA	889	30	2.9	108554	9	AC005226	Homo sapi
817	31	3.0	173957	2	AC091671	Papio cyn	890	30	2.9	114983	2	AC0933209	Human DNA
818	31	3.0	175835	2	AC024074	Homo sapi	891	30	2.9	115932	9	HS1107	Homo sapi
819	31	3.0	176388	2	AC025273	Homo sapi	892	30	2.9	117031	9	AC008973	Homo sapi
c 820	31	3.0	176407	2	AP003040	Homo sapi	893	30	2.9	119942	2	AC024582	Homo sapi
c 821	31	3.0	178045	2	AC092719	Homo sapi	c 894	30	2.9	119942	2	AC024582	Homo sapi
c 822	31	3.0	178325	2	AC027601	Homo sapi	c 895	30	2.9	120997	9	AC013449	Homo sapi
c 823	31	3.0	178434	2	AC073046	Homo sapi	896	30	2.9	121208	9	AL451073	Human DNA
824	31	3.0	178904	2	AL354989	Homo sapi	897	30	2.9	122103	9	HS291J10	Human DNA
c 825	31	3.0	179974	2	AC021704	Homo sapi	898	30	2.9	124281	9	AL136101	Human DNA
c 826	31	3.0	180044	2	AC026942	Homo sapi	899	30	2.9	124437	9	AL139125	Homo sapi
c 827	31	3.0	180942	33	AC019194	Homo sapi	900	30	2.9	131692	2	AC079927	Homo sapi
828	31	3.0	181829	9	AC023425	Homo sapi	901	30	2.9	131841	9	HSJ881L22	Human DNA
c 829	31	3.0	183451	2	AC024348	Homo sapi	c 902	30	2.9	134594	2	AC083822	Homo sapi
c 830	31	3.0	183813	2	AC012204	Homo sapi	c 903	30	2.9	135305	2	AC090427	Homo sapi
831	31	3.0	184637	2	AC012043	Human DNA	c 904	30	2.9	136129	2	AC092483	Homo sapi
832	31	3.0	185229	9	AL157762	Human DNA	c 905	30	2.9	136155	2	AC069346	Homo sapi
c 833	31	3.0	185401	2	AC069068	Homo sapi	906	30	2.9	139488	9	AC003043	Homo sapi
c 834	31	3.0	185804	2	AC017095	Homo sapi	907	30	2.9	139836	2	AC090052	Homo sapi
835	31	3.0	187072	2	AL161733	Homo sapi	908	30	2.9	140026	9	AC005840	Homo sapi
c 836	31	3.0	188527	9	AC084357	Homo sapi	909	30	2.9	141049	2	AL354662	Homo sapi
c 837	31	3.0	189608	2	AC018921	Homo sapi	c 910	30	2.9	141597	2	AC008455	Homo sapi
c 838	31	3.0	190579	2	AC018736	Homo sapi	c 911	30	2.9	142344	2	AC093223	Homo sapi
c 839	31	3.0	191905	2	AL512770	Homo sapi	c 912	30	2.9	145340	9	AL136525	Homo sapi
c 840	31	3.0	191942	9	AC079951	Homo sapi	913	30	2.9	145833	2	AC026539	Homo sapi
c 841	31	3.0	192768	9	AC016831	Human DNA	c 914	30	2.9	146487	2	AC022243	Homo sapi
c 842	31	3.0	196000	9	AL354749	Human DNA	915	30	2.9	149901	2	AC022243	Homo sapi
c 843	31	3.0	196327	2	AC092429	Homo sapi	916	30	2.9	149913	2	AC010330	Homo sapi
c 844	31	3.0	197791	2	AC016153	Homo sapi	c 917	30	2.9	149913	2	AC010330	Homo sapi
845	31	3.0	199380	2	AC048379	Homo sapi	918	30	2.9	151845	2	AC008128	Homo sapi
846	31	3.0	200000	2	AC004618	Homo sapi	c 919	30	2.9	153087	2	AC017072	Homo sapi
c 847	31	3.0	200536	9	AL356093	Human DNA	c 920	30	2.9	153716	2	AC069187	Homo sapi
c 848	31	3.0	200594	9	AC005632	Homo sapi	921	30	2.9	154270	2	AC023391	Homo sapi
c 849	31	3.0	201794	9	CNS018OW	Human chr	922	30	2.9	154745	2	AC010860	Homo sapi
850	31	3.0	204272	2	AC007215	Homo sapi	923	30	2.9	154960	2	AC011948	Homo sapi
c 851	31	3.0	210000	2	AC004555	Homo sapi	924	30	2.9	155459	2	AC016619	Homo sapi
c 852	31	3.0	215727	2	AL358795	Homo sapi	c 925	30	2.9	155769	2	AC073250	Homo sapi
c 853	31	3.0	220128	2	AL139407	Homo sapi	926	30	2.9	156357	9	AC011456	Homo sapi
c 854	31	3.0	236680	2	AC024120	Homo sapi	c 927	30	2.9	156390	2	AL358432	Homo sapi
c 855	31	3.0	237044	2	AP000926	Homo sapi	928	30	2.9	157058	2	CNS05TBY	Homo sapi
c 856	31	3.0	247473	2	AC011137	Homo sapi	929	30	2.9	157186	2	AL592076	Homo sapi
c 857	31	3.0	340000	9	AP001725	Homo sapi	c 930	30	2.9	160209	9	AC012324	Homo sapi
c 858	30	2.9	1637	9	AF349314	Sequence	c 931	30	2.9	161198	2	AC015867	Homo sapi
c 859	30	2.9	1762	6	AX147248	Sequence	c 932	30	2.9	161264	9	AC007011	Homo sapi
860	30	2.9	1804	9	AK022325	Homo sapi	c 933	30	2.9	162486	2	AC015643	Homo sapi
c 861	30	2.9	1916	9	AK023510	Homo sapi	c 934	30	2.9	162575	9	AL157829	Human DNA
862	30	2.9	1994	9	BC008304	Homo sapi	c 935	30	2.9	162962	2	AC021790	Homo sapi
c 863	30	2.9	12845	2	AC092308	Homo sapi	c 936	30	2.9	162962	2	AC021790	Homo sapi
c 864	30	2.9	14402	2	AC009168	Homo sapi	c 937	30	2.9	164241	2	AL157935	Homo sapi
865	30	2.9	17982	2	AC009168	Homo sapi	c 938	30	2.9	164628	2	AC073646	Homo sapi
c 866	30	2.9	17982	2	AC093239	Homo sapi	c 939	30	2.9	164662	2	AC024560	Homo sapi
c 867	30	2.9	23179	9	AC007487	Homo sapi	940	30	2.9	164733	2	AC015602	Homo sapi
c 868	30	2.9	29979	9	AC000122	Homo sapi	941	30	2.9	164884	2	AC020909	Homo sapi
869	30	2.9	32260	2	AC093237	Homo sapi	942	30	2.9	165868	2	AC021719	Homo sapi
c 870	30	2.9	32260	2	AC093237	Homo sapi	943	30	2.9	166264	2	AL390864	Homo sapi
871	30	2.9	36236	9	AL589656	Human DNA	944	30	2.9	166449	2	AC093012	Homo sapi
872	30	2.9	39383	9	AC022149	Homo sapi	c 945	30	2.9	166915	2	AL596093	Homo sapi
c 873	30	2.9	43390	9	AC007191	Homo sapi	c 946	30	2.9	166961	2	AL353137	Homo sapi

947	30	2.9	167074	2	AC026643
948	30	2.9	167357	9	CNS01DWY
949	30	2.9	167408	2	AC023020
c 950	30	2.9	168149	2	AL158843
951	30	2.9	168230	2	AC010624
c 952	30	2.9	168230	2	AC010624
953	30	2.9	169323	2	AL161906
954	30	2.9	169343	9	AC016955
955	30	2.9	169623	2	AC055764
956	30	2.9	169845	2	AC019054
957	30	2.9	170114	2	AC016143
958	30	2.9	170470	2	AC022696
959	30	2.9	170655	2	AP001012
c 960	30	2.9	170655	2	AP001012
961	30	2.9	170954	2	AL356740
c 962	30	2.9	171189	2	AC008966
c 963	30	2.9	171788	9	AC073089
964	30	2.9	171922	2	AC091045
965	30	2.9	172601	2	AC021455
c 966	30	2.9	173149	2	AL391218
c 967	30	2.9	173213	2	AC066692
968	30	2.9	173452	9	AC003013
969	30	2.9	173575	2	AC016932
970	30	2.9	173646	2	AC016542
c 971	30	2.9	174466	2	AC068768
c 972	30	2.9	174476	2	AC027687
c 973	30	2.9	174480	2	AC026069
974	30	2.9	174579	2	AC016666
975	30	2.9	175271	2	AC090698
c 976	30	2.9	175670	2	AL358353
977	30	2.9	175789	2	AP000425
c 978	30	2.9	175999	2	AC021154
979	30	2.9	176391	9	AL365400
980	30	2.9	176571	2	AC027111
981	30	2.9	177066	2	AC069257
c 982	30	2.9	177289	2	AC090176
c 983	30	2.9	178246	2	AL591506
984	30	2.9	179134	2	AC022817
985	30	2.9	180486	2	AC007621
986	30	2.9	180520	2	AC019152
c 987	30	2.9	180522	2	CNS01DSH
988	30	2.9	182018	2	AC015464
c 989	30	2.9	182047	2	AC013462
c 990	30	2.9	182214	2	AC067869
991	30	2.9	182414	2	AC055877
992	30	2.9	182747	2	AL512330
993	30	2.9	183842	2	AC080084
994	30	2.9	184476	2	AL138750
995	30	2.9	184853	2	AC020723
996	30	2.9	186607	2	AC092981
997	30	2.9	187760	2	AC092755
c 998	30	2.9	188076	2	AC023629
c 999	30	2.9	188249	2	AL450023
1000	30	2.9	188346	2	AC010189

ALIGNMENTS

RESULT	1
AX107720	
LOCUS	AX107720 1041 bp DNA
DEFINITION	Sequence 9 from Patent WO0123579.
ACCESSION	AX107720
VERSION	AX107720.1 GI:13923201
KEYWORDS	.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
	1 (bases 1 to 1041)
	Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.

AC026643	Homo sapi
AL138976	Human chr
AC023020	Homo sapi
AL158843	Homo sapi
AC010624	Homo sapi
AC010624	Homo sapi
AL161906	Homo sapi
AL161906	Homo sapi
AC016955	Homo sapi
AC055764	Homo sapi
AC019054	Homo sapi
AC016143	Homo sapi
AC022696	Homo sapi
AP001012	Homo sapi
AP001012	Homo sapi
AL356740	Homo sapi
AC008966	Homo sapi
AC073089	Homo sapi
AC091045	Homo sapi
AC021455	Homo sapi
AL391218	Homo sapi
AC066692	Homo sapi
AC003013	Human PAC
AC016932	Homo sapi
AC016542	Homo sapi
AC068768	Homo sapi
AC027687	Homo sapi
AC026069	Homo sapi
AC016666	Homo sapi
AC090698	Homo sapi
AL358353	Homo sapi
AP000425	Homo sapi
AC021154	Homo sapi
AL365400	Human DNA
AC027111	Homo sapi
AC069257	Homo sapi
AC090176	Homo sapi
AL591506	Homo sapi
AC022817	Homo sapi
AC007621	Homo sapi
AC019152	Homo sapi
AL121784	Homo sapi
AC015464	Homo sapi
AC013462	Homo sapi
AC067869	Homo sapi
AC055877	Homo sapi
AL512330	Homo sapi
AC080084	Homo sapi
AL138750	Homo sapi
AC020723	Homo sapi
AC092981	Homo sapi
AC092755	Homo sapi
AC023629	Homo sapi
AL450023	Homo sapi
AC010189	Homo sapi

TITLE	Human kinase proteins and polynucleotides encoding the same
JOURNAL	Patent: WO 0123579-A 9 05-APR-2001;
FEATURES	Lexicon Genetics Incorporated (US)
Source	Location/Qualifiers
	1..1041
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
BASE COUNT	344 a 190 c 228 g 279 t
ORIGIN	
Query Match	100.0%; Score 1041; DB 6; Length 1041;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 1041: Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 atggaagagatgaaaaattagctaaagactggagaagggtcttattgggggtgtattc
Db	1 ATCGAAAAAGTATGAAAAATTAGCTAAGACTGGAGAAGGGTCTATGGGGTGTATTCAA
QY	61 tgcagaaacaaaacctctggacaagtagctgttaaaaaatttgggaatctgaagat
Db	61 TGCAGAAACAAAACCTCTGCACAAGTAGTAGCTCTTAAAAAATTGTGGAATCTGAAGAT
QY	121 gactctgtgttaagaaaaatagcactaagagaaatcacgtatgttgaagcaattaaacat
Db	121 GATCCTGTTGTTAAGAAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAAACAT
QY	181 ccaaatcttgtgaacctcatcacttttaaatgagctggaaagaaacccaatggagttgctgat
Db	181 CCAAAATCTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAATGGAGTTGCTGAT
QY	241 gaatactgtgatcatacacacttttaaatgagctggaaagaaacccaatggagttgctgat
Db	241 GAATACGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT
QY	301 ggagtgatcaaaaagcgattattatggcacaacacttcaagctcttaattctgtcatatacat
Db	301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTCTGTCATATACAT
QY	361 aactgtattcacagagatataaaaacctgaaataattcttaataactaagcaaggaataatc
Db	361 AACTGTATTCACAGAGATATAAAACCTGAAAAATATTCTAATAAGCAAGGAATAATC
QY	421 aagatttgtgacttcgggttgacaaaacttctgatccaggagatgcctacacccgattat
Db	421 AAGATTGTGACTTCGGGTTTGCACAAATCTGTATCCAGGAGATGCCTACACCGATTAT
QY	481 gtagctacgagatggtacogagctcctgaacttcttgtgtggagatactcagtaggtctct
Db	481 GTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTGTGGGAGATACTCAGTATGGTTCT
QY	541 tcagtcgatatatggctattgtgtgtttttgtcagagctcctgacagggccagccactg
Db	541 TCAGTCGATATATGGGCTATTGGTTGTGTTTTCAGAGCTCTTGACAGGCTCCAGCCACTG
QY	601 tggcctggaaaaatcagatgtggaccaacttttatctgtataatcagaaacactagtagagacg
Db	601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGTAGAGACG
QY	661 ggggttcgccatgtttgaccaggctgggtctcgaaactcttgacgtcaagtgtaccactgccc
Db	661 GGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTCTTGACGTCAAGTGATCCACCTGCC
QY	721 gtagcctctcaaaagtgcctggaattacagggaaaaattaatcccaagacatcaatcatttt
Db	721 GTAGCCTCTCAAAGTGCCTGGAATTACAGGAAAAATTATCCCAAGACATCAATCAATCTTT
QY	781 aaaagtaacgggtttttccatggcatcagtagtataccigagccagagacatggaactctt
Db	781 AAAAGTAAACGGGTTTTTCCATGGCATCAGTATACCTCAGCCAGAGACATGGAAACTCTT
QY	841 gaggaagagttctcagatgttcatcctgtggctctgaacttcatgaagggtgtctgaag

Fri May 3 11:02:09 2002

Db	841	GAGGAAAAATTCTCAGATGTTTCATCCTGTGGCTCTGAACCTTCATGAAGGGTGCTCTGAAG	900
QY	901	atgaatccagatgacagattaacacctgttcccaactcctggagagctcctacttggattct	960
Db	901	ATGAATCCAGATGACAGATTAACTGTTCCCAACTCCTGGAGAGTCCTACTTTGATTCT	960
QY	961	tttcaagaggcccaaatataaagaaaaagcacgtaatgaaggaaagaacacagacgcgcaa	1020
Db	961	TTTCAAGAGGCCCAAAATTAAAGAAAGCAGCGTAATGAAGGAAGAAACAGAACGCCAA	1020
QY	1021	caggtacttcgcgtcaaaaagt	1041
Db	1021	CAGGTACTTCCGCTCAAAAGT	1041
RESULT 2			
AX107714	AX107714	1068 bp	DNA
LOCUS	Sequence 3	from Patent	WO0123579.
DEFINITION	Sequence 3	from Patent	WO0123579.
ACCESSION	AX107714		
VERSION	AX107714.1	GI:13923198	
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 1068)		
AUTHORS	Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.		
TITLE	Human kinase proteins and polynucleotides encoding the same		
JOURNAL	Patent: WO 0123579-A 3 05-APR-2001;		
FEATURES	Lexicon Genetics Incorporated (US)		
source	Location/Qualifiers		
	1..1068		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"	287 t	
BASE COUNT	344 a	197 c	240 g
ORIGIN			
Query Match 98.5%; Score 1025; DB 6; Length 1068;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 1025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	atggaaaagtatgaaaaattagctaagactggagaagggtcttatgggggtgtattcaaa	60
Db	1	ATGGAAAAGTATGAAAAATTAGCTAAGACTGGAGAAGGGTCTTATGGGGTTGTATTCAA	60
QY	61	tcagaaaaaacctctggacaagttagtgcgtgttaaaaaatttgggaactctgaagat	120
Db	61	TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTGTAAAAAATTGTGGAAATCTGAAGAT	120
QY	121	gacctgtgttaagaaaatagcactaagagaaaatacgtatgttgaagcaattaaaaacat	180
Db	121	GATCCTGTTGTTAAGAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAAACAT	180
QY	181	ccaaatcttggaaacctcatcgagggttcaggagaaaaaggaaaatgcatttagtttt	240
Db	181	CCAAATCTTGTAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAATGCATTAGTTTIT	240
QY	241	gaatactgtgatcacatacacacttttaaatgagctggaaagaaacccaaatggagtgcgtg	300
Db	241	GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCCAATGGAGTGTGTAT	300
QY	301	ggagtgatcaaaagcgtattatggcaaacacttcaagctcttaattctgtcatatacat	360
Db	301	GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCTATATACAT	360
QY	361	aactgtattcacagagatataaaaacctgaaaataattcttaactaagcaaggaataatc	420
Db	361	AACTGTATTCACAGAGATATAAAACCTGAAAATATTCTTAATAACTAAGCAAGGAATAATC	420
QY	421	aagatttgtgacttcgggtttgcacaaaattctgattccaggagatgcctacacccgattat	480

Db	421	AAGATTGTGACTCGGGTTGCACAAATTTCTGATTCCAGGAGATGCCTACACCGATTAT	480
QY	481	gtagtcacgagatggttaccgagctcctgaactcttltgtgggagatactcagtatggttct	540
Db	481	GTAGCTACGAGATGTTACCGAGCTCCTGAACCTTCTTGTGGGAGATACTCAGTATGGTTCT	540
QY	541	tcagtcgatatatgggctattgggttltgtttttgcagagctcctgacagggccagccactg	600
Db	541	TCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAGCCACTG	600
QY	601	tggcctggaaaaatcagatgtggaccaactttatctgatataatcagaaacactagtagacg	660
Db	601	TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATATCAGAACACTAGTAGAGACG	660
QY	661	gggtttcgccatgttgaccagcgtggtctctcgaactcttgacgtcaagtcatccacctgcc	720
Db	661	GGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTCTTGACCTCAAGTGATCCACCTGCC	720
QY	721	gtagcctctcaaaagtgtggaattacaggaaaaataatcccaagacatcaatcaatcttt	780
Db	721	GTAGCCTCTCAAAAGTGTGGAATTACAGGAAAAATTAATCCCAAGACATCAATCAATCTTT	780
QY	781	aaaagtaacgggtttttccatggcatcacgtatatacctgagccagagacacatggaactctt	840
Db	781	AAAAGTAAACGGGTTTTTCCATGGCATCAGTATACCTTGAGCCAGAGACATGGAAACTCTT	840
QY	841	gaggaaaaagttctcagatgttccatcctgtggctctcgaacttcatgaagggtgtctgaag	900
Db	841	GAGGAAAAGTTCICAGATGTTTCATCCTGTGGCTCTGAACCTTCATGAAGGGGTGTCTGAAG	900
QY	901	atgaatccagatgacagattaaacctgttcccaactcctggagagctcctacttctgattct	960
Db	901	ATGAATCCAGATGACAGATTAACCTGTTCCCAACTCCTGGAGAGCTCTACTTTGATTCT	960
QY	961	tttcaagaggcccaaatataaagaaaaagcacgtaatgaaggaaagaaacagagagcga	1020
Db	961	TTTCAAGAGGCCCAAAATTAAAGAAAAAGCACGCTAAATGAAGGAAGAAACAGAGACGCCAA	1020
QY	1021	caggt 1025	
Db	1021	CAGGT 1025	
RESULT 3			
AX107722	AX107722	945 bp	DNA
LOCUS	Sequence 11	from Patent	WO0123579.
DEFINITION	Sequence 11	from Patent	WO0123579.
ACCESSION	AX107722		
VERSION	AX107722.1	GI:13923202	
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 945)		
AUTHORS	Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.		
TITLE	Human kinase proteins and polynucleotides encoding the same		
JOURNAL	Patent: WO 0123579-A 11 05-APR-2001;		
FEATURES	Lexicon Genetics Incorporated (US)		
source	Location/Qualifiers		
	1..945		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"	255 t	
BASE COUNT	323 a	165 c	202 g
ORIGIN			
Query Match 62.6%; Score 652; DB 6; Length 945;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 652; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			

QY 1 atggaaaagtatgaaaaattagctaaagctggaagggtcttattatgggggtgtattcaaa 60
|||||
Db 1 ATGGAAAAGTATGAAAAATTAGCTAAGACTGGAGAAGGGCTTATGGGGTTGTATTCAAA 60

QY 61 tgcagaaaacaaacctctggacaagtagtagctgttaaaaaatttgggaatctgaagat 120
|||||
Db 61 TGCAGAAACAAACCTCTGGACAAGTAGCTGTATAAAAATTGTGGAATCTGAAGAT 120

QY 121 gatcctgtgttaagaaaaatagcactaagagaaatcacgtatgtgaagcaattaaaaacat 180
|||||
Db 121 GATCCTGTGTTAAGAAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAAAAACAT 180

QY 181 ccaaatcttgtgaacctcatcgaggtgttcaggagaaaaagaaaaatgcatttagtttt 240
|||||
Db 181 CCAAAATCTGTGAACCTCATCGAGGTGTTCAGGAGAAAAAGCAATTTAGTTT 240

QY 241 gaatactgtgatcatcacacttttaaatgagctggaagaaaccccaaatggagttgctgat 300
|||||
Db 241 GAACTACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATCGAGTTCTTGAT 300

QY 301 ggaagtgatcaaaagcgtattatggcaaacacacttcaagctcttaatttctgtcatatacat 360
|||||
Db 301 GGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTCAATACAT 360

QY 361 aactgtattcacagagatatataaacctgaaaaatattctaataactaaaggaataaatc 420
|||||
Db 361 AACTGTATTACAGAGATATAAAACCTGAAATATCTAATPAACTAAGCAAGGAATAATC 420

QY 421 aagatttgtgacttcgggtttgcacaaattctgattccaggagatgcctcacaccgattat 480
|||||
Db 421 AAGATTTGTGACTTCGGGTTTGCACAAATTCGATTCCAGGAGATGCCATACACCGATTAT 480

QY 481 gtagctacgagatggtaccgagctcctgaacttcttgtgggagatactcagtatggttct 540
|||||
Db 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTTGTGGGAGATACATCAGTATGTTCT 540

QY 541 tcagtcgatatatgggctattgggttgtttttgcagagctcctgacagggccagccactg 600
|||||
Db 541 TCAGTCGATATATGGCTATTGGTTGTGTTTTCGAGAGCTCCTGACAGGCCAGCCACTG 600

QY 601 tggcctgaaaaatcagatgtggaccaactttatctgataatcagaaacactag 652
|||||
Db 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAAACACTAG 652

RESULT 4
AX107716
LOCUS AX107716 972 bp DNA PAT 30-APR-2001
DEFINITION Sequence 5 from Patent WO0123579.
ACCESSION AX107716
VERSION AX107716.1 GI:13923199
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 972)
AUTHORS Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.
TITLE Human Kinase proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0123579-A 5 05-APR-2001;
Lexicon Genetics Incorporated (US)
FEATURES Location/Qualifiers
source 1..972
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 323 a 172 c 214 g 263 t
ORIGIN

Query Match 62.6%; Score 652; DB 6; Length 972;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 652; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atggaaaagtatgaaaaattagctaaagctggaagggtcttattatgggggtgtattcaaa 60
|||||
Db 1 ATGGAAAAGTATGAAAAATTAGCTAAGACTGGAGAAGGGTCTTATGGGGTTGTATTCAAA 60

QY 61 tgcagaaaacaaacctctggacaagtagtagctgttaaaaaatttgtggaatctgaagat 120
|||||
Db 61 TGCAGAAACAAACCTCTGGACAAGTAGCTGTATAAAAATTGTGGAATCTGAAGAT 120

QY 121 gatcctgtgttaagaaaaatagcactaagagaaatcacgtatgtgaagcaattaaaaacat 180
|||||
Db 121 GATCCTGTGTTAAGAAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAAAAACAT 180

QY 181 ccaaatcttgtgaacctcatcgaggtgttcaggagaaaaagaaaaatgcatttagtttt 240
|||||
Db 181 CCAAAATCTGTGAACCTCATCGAGGTGTTCAGGAGAAAAAGCAATTTAGTTT 240

QY 241 gaatactgtgatcatcacacttttaaatgagctggaagaaaccccaaatggagttgctgat 300
|||||
Db 241 GAACTACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGAT 300

QY 301 ggaagtgatcaaaagcgtattatggcaaacacacttcaagctcttaatttctgtcatatacat 360
|||||
Db 301 GGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTCAATACAT 360

QY 361 aactgtattcacagagatatataaacctgaaaaatattctaataactaaaggaataaatc 420
|||||
Db 361 AACTGTATTACAGAGATATAAAACCTGAAATATTTCTAATAACTAAGCAAGGAATAATC 420

QY 421 aagatttgtgacttcgggtttgcacaaattctgattccaggagatgcctcacaccgattat 480
|||||
Db 421 AAGATTTGTGACTTCGGGTTTGCACAAATTCGATTCCAGGAGATGCCATACACCGATTAT 480

QY 481 gtagctacgagatggtaccgagctcctgaacttcttgtgggagatactcagtatggttct 540
|||||
Db 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTTGTGGGAGATACTCAGTATGTTCT 540

QY 541 tcagtcgatatatgggctattgggttgtttttgcagagctcctgacagggccagccactg 600
|||||
Db 541 TCAGTCGATATATGGCTATTGGTTGTGTTTTCGAGAGCTCCTGACAGGCCAGCCACTG 600

QY 601 tggcctgaaaaatcagatgtggaccaactttatctgataatcagaaacactag 652
|||||
Db 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAAACACTAG 652

RESULT 5
AX107724
LOCUS AX107724 1819 bp DNA PAT 30-APR-2001
DEFINITION Sequence 13 from Patent WO0123579.
ACCESSION AX107724
VERSION AX107724.1 GI:13923203
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1819)
AUTHORS Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.
TITLE Human Kinase proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0123579-A 13 05-APR-2001;
Lexicon Genetics Incorporated (US)
FEATURES Location/Qualifiers
source 1..1819
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 539 a 368 c 406 g 505 t 1 others
ORIGIN

Query Match 54.9%; Score 572; DB 6; Length 1819;
Best Local Similarity 100.0%; Pred. No. 2.1e-306;

Matches 572; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy 454 attcaggagatccttacacgattatgtagctacgagatggtaccgagctcctgaactt 513	
Db 897 ATTCCAGGAGATGCTACACCGATTATGTAGCTACGAGATGGTACCGAGCTCCTGAACIT 956	
Qy 514 cttgtggagatactcagtatggttcttcagtcgatatatggctattggtggttttt 573	
Db 957 CTTGTGGGAGATACTCAGTATGTTCTTCAGTCGATATATGGGCTATTGGTTGTGTTTT 1016	
Qy 574 gcagagctcctgacagggccagccactgtggcctggaaatcagatgtggaccactttat 633	
Db 1017 GCAGAGCTCCTGACAGGCCAGCCACTGTGGCCTGGAAATCAGATGTGGACCACTTTAT 1076	
Qy 634 ctgataatcagaacactagtagagacggggttccgcatgttgaccagggctggtctcgaa 693	
Db 1077 CTGATAATCAGAACACTAGTAGACGGGGTTTCGCCATGTTGACCAAGCTGCTCGAA 1136	
Qy 694 ctcttgacgtcaagtgtatccacctgocgtagcctctcctcaagtgtggaattacagga 753	
Db 1137 CTCTTGACGTCAAGTGATCCACCTGCCGTAGCCTCTCAAAAGTGCTGGAATTACAGGAAA 1196	
Qy 754 ttaatcccaagacatcaatcaatctttaaaagtaacgggttttccatggcatcagtata 813	
Db 1197 TTAATCCCAAGACATCAATCAATCTTTAAAGTAACGGGTTTTCATGTCATCAGTATA 1256	
Qy 814 cctgagccagagacatggaactcttgaggaaagtcttcagatgttcactcctgtggt 873	
Db 1257 CCTGAGCCAGAGACATGGAAACTCTTGAGGAAAGTTCTCAGATGTTCACTCTGTGGCT 1316	
Qy 874 ctgaacttcgatgaagggtgtctgaaagtgaatccagatgacagattaacctgttcccaa 933	
Db 1317 CTGAACCTTCATGAAGGGGTGTCTGAAGATGAATCCAGATGACAGATTAACCTGTTCCCAA 1376	
Qy 934 ctctggagagctcctacttcttgattctttcaagagggcccaattaaaaagaaagcacgt 993	
Db 1377 CTCCTGGAGAGCTCCTACTTTTGATTCTTTCAAGAGGCCCAAAATTAAGAAAGAACACGT 1436	
Qy 994 aatgaaggagaacacagagacgccaacaggt 1025	
Db 1437 AATGAAGGAAGAAACAGAGACGCCCAACAGGT 1468	
RESULT 6	
AX107712 AX107712 561 bp DNA PAT 30-APR-2001	
LOCUS Sequence 1 from Patent WO0123579.	
DEFINITION AX107712	
ACCESSION AX107712	
VERSION AX107712.1 GI:13923197	
KEYWORDS	
SOURCE human.	
ORGANISM Homo sapiens	
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS 1 (bases 1 to 561)	
Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.	
TITLE Human kinase proteins and polynucleotides encoding the same	
JOURNAL Patent: WO 0123579-A 1 05-APR-2001;	
Lexicon Genetics Incorporated (US)	
FEATURES Location/Qualifiers	
source 1..561	
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
BASE COUNT 189 a 87 c 112 g 173 t	
ORIGIN	
Query Match 43.6%; Score 454; DB 6; Length 561;	
Best Local Similarity 100.0%; Pred. No. 8.2e-241;	
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy 1 atggaaaagtatgaaaaattagctaagactggagagggtcttattggttatttcaaa 60	

Db 1 ATGGAAAAGTATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTTGTTATTCAAA 60	
Qy 51 tgcagaaaacaaacctctggacaagtagtagctgtttaaaaaatttgtggaatctgaagat 120	
Db 61 TGCAGAAAACAAACCTCTGGACAAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT 120	
Qy 121 gatcctgttgtttaagaaaaatagcactaagagaaaatcgtatgttggaagcaattaaaaacat 180	
Db 121 GATCCTGTGTGTTAAGAAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTTAAACAT 180	
Qy 181 ccaaatcttgtgaacctcatcgaggtgttcaggagagaaaaaggaaaaatgcatttagtttt 240	
Db 181 CCAAAATCTTGTGAACCTCATCGAGGTGTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 240	
Qy 241 gaatactgtgatcatcacactttttaaatgagctggaaagaaaccccaaatggagttgctgat 300	
Db 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300	
Qy 301 ggagtgatcaaaagcgtatttatggcaaacaccticcaagctcttaatttctgtcatatacat 360	
Db 301 GGAGTGATCAAAAAGCGTATTATGGCAACACCTTCAAGCTCTTAAATTTCTGTCATATACAT 360	
Qy 361 aactgtattcacagagatataaaaaacctgaaaaatttctaataactaagcaaggaataatc 420	
Db 361 AACTGTATTACAGAGATATAAAAACCTGAAAAATATTCTAATACTAAGCAAGGAATAATC 420	
Qy 421 aagatttgtgactcggggttgcacaaaattctga 454	
Db 421 AAGATTGTGACTTCGGGTTTGCACAAATCTGA 454	
RESULT 7	
AX107718 AX107718 594 bp DNA PAT 30-APR-2001	
LOCUS Sequence 7 from Patent WO0123579.	
DEFINITION AX107718	
ACCESSION AX107718	
VERSION AX107718.1 GI:13923200	
KEYWORDS	
SOURCE human.	
ORGANISM Homo sapiens	
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS 1 (bases 1 to 594)	
Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.	
TITLE Human kinase proteins and polynucleotides encoding the same	
JOURNAL Patent: WO 0123579-A 7 05-APR-2001;	
Lexicon Genetics Incorporated (US)	
FEATURES Location/Qualifiers	
source 1..594	
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
BASE COUNT 203 a 94 c 115 g 182 t	
ORIGIN	
Query Match 43.6%; Score 454; DB 6; Length 594;	
Best Local Similarity 100.0%; Pred. No. 8.2e-241;	
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy 1 atggaaaagtatgaaaaattagctaagactggagagggtcttattggttatttcaaa 60	
Db 1 ATGGAAAAGTATGAAAAATTTAGCTAAGACTGGAGAGGGTCTTATGGGGTTGTTATTCAAA 60	
Qy 61 tgcagaaaacaaacctctggacaagtagtagctgtttaaaaaatttgtggaatctgaagat 120	
Db 61 TGCAGAAAACAAACCTCTCGACAAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT 120	
Qy 121 gatcctgttgtttaagaaaaatagcactaagagaaaatcgtatgttggaatctgaagat 180	
Db 121 GATCCTGTGTGTTAAGAAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTTAAACAT 180	

QY 181 ccaaatcttgtgaacctcatcgaggtgttcaggagaaaaaggaaaaatgcatttagtttt 240
|||||
Db 181 CCAAATCTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAGGAAATGCATTTAGTTTTT 240
QY 241 gaatactgtgatcatcacacttttaaatgagctggaaagaaaccccaaatggagttgctgat 300
|||||
Db 241 GAATACGTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCCAATGGAGTGTCTGAT 300
QY 301 ggagtgatcaaaaagcgtattatggaacacacttcaagctcttaatttctgtcatatacat 360
|||||
Db 301 GGAGTGATCAAAAAGCGTATTATTGGCAACACTTCAAGCTCTTTAATTTCTGTATATACAT 360
QY 361 aactgtattcacagagatataaaacctgaaaaattcttaataactaaagcaaggaataatc 420
|||||
Db 361 AACTGTATTTCACAGAGATATAAAACCTGAAAATATTCTAATACTAAGCAAGGAATAATC 420
QY 421 aagatttgtgacttcgggtttgacacaaattctga 454
|||||
Db 421 AAGATTGTGACTTCGGGTTTGCACAAATTTCTGA 454

RESULT 8
AX166534
LOCUS AX166534 1083 bp DNA PAT 22-JUN-2001
DEFINITION Sequence 25 from Patent WO0138503.
ACCESSION AX166534
VERSION AX166534.1 GI:14546879
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1083)
AUTHORS Plowman,G.D., Whyte,D., Manning,G.S., Sudarsanam,S.S., Martinez,R.,
Flanagan,P. and Clary,D.S.
TITLE Novel human protein kinases and protein kinase-like enzymes
JOURNAL Patent: WO 0138503-A 25 31-MAY-2001;
Sugen, Inc. (US)
FEATURES
source Location/Qualifiers
1..1083
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 366 a 199 c 225 g 293 t
ORIGIN

Query Match 43.6%; Score 454; DB 6; Length 1083;
Best Local Similarity 100.0%; Pred. No. 8.3e-241;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atggaaaaagtataaaaaattagctaagactggagaagggtcttatgggttgattcaaa 60
|||||
Db 1 ATGGAAAAGTATGAAAAATTAGCTAAGACTGGAGAAGGGTCTTATGGGTTGTATTCAA 60
QY 61 tgcagaaaaaaaacctctggacaagtagtagctgttaaaaaaatttgtggaatctgaagat 120
|||||
Db 61 TGCAGAAAACAAAACCTCTGGACAAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT 120
QY 121 gatcctgttgaagaaaaatagcacLaagagaatacgtatgttgaagcaattaaaaacat 180
|||||
Db 121 GATCCTGTTGTTAAGAAAAATAGCATAAGAGAAATACGTATGTTGAAGCAATTTAAACAT 180
QY 181 ccaaatcttgtgaacctcatcgaggtgttcaggagaaaaaggaaaaatgcatttagtttt 240
|||||
Db 181 CCAAATCTGTGAACCTCATCGAGGTGTTCAGGAGAAAAGGAAATGCATTTAGTTTTT 240
QY 241 gaatactgtgatcatcacacttttaaatgagctggaaagaaaccccaaatggagttgctgat 300
|||||
Db 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCCAATGGAGTTGCTGAT 300
QY 301 ggagtgatcaaaaagcgtattatggaacacacttcaagctcttaatttctgtcatatacat 360
|||||
Db 301 GGAGTGATCAAAAAGCGTATTATTGGCAACACTTCAAGCTCTTTAATTTCTGCATATACAT 360

QY 361 aactgtattcacagagatataaaacctgaaaaattcttaataactaaagcaaggaataatc 420
|||||
Db 361 AACTGTATTTCACAGAGATATAAAACCTGAAAATATTCTAATACTAAGCAAGGAATAATC 420
QY 421 aagatttgtgacttcgggtttgacacaaattctga 454
|||||
Db 421 AAGATTGTGACTTCGGGTTTGCACAAATTTCTGA 454

RESULT 9
AX056404
LOCUS AX056404 911 bp DNA PAT 13-JAN-2001
DEFINITION Sequence 48 from Patent WO0073469.
ACCESSION AX056404
VERSION AX056404.1 GI:12229111
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 911)
AUTHORS Plowman,G.D., Martinez,R., Whyte,D. and Sudersanam,S.
TITLE Protein kinases
JOURNAL Patent: WO 0073469-A 48 07-DEC-2000;
Sugen, Inc. (US)
FEATURES
source Location/Qualifiers
1..911
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 293 a 182 c 178 g 258 t
ORIGIN

Query Match 35.4%; Score 368; DB 6; Length 911;
Best Local Similarity 100.0%; Pred. No. 5.3e-193;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 285 aaatggagttgctgatggagtgatcaaaagcgtattatggcaaacacttcaagctcttaa 344
|||||
Db 1 AAATGGAGTTGCTGATGGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAA 60
QY 345 ttctgtcatatacatataactgtatttcacagagatataaaacctgaaaaatattctaatac 404
|||||
Db 61 TTTCTGTCATATACATAACTGTATTTCACAGAGATATAAAACCTGAAAAATATTCTAATAAC 120
QY 405 taagcaaggaaataatcaagatttctgacttcgggtttgcacaaattctgattccaggaga 464
|||||
Db 121 TAAGCAAGGAATAATCAAGATTTGTGACTTCGGGTTTGCACAAATTTCTGATTCAGGAGA 180
QY 465 tgcctacaccgattatgtagctacgagatggtaccgagctcctgaacttctgtgggaga 524
|||||
Db 181 TGCTACACCGATTATGTAGTACGAGATGGTACCGAGCTCTGAACTTCTTGTGGGAGA 240
QY 525 tactcagtgatggttcttcagtcgatataatgggtattlgggtgtgttttcagagagtcct 584
|||||
Db 241 TACTCAGTATGGTTCTTCAGTCGATATATGGGCTATTGGTGTGTTTTGCAGAGCTCCT 300
QY 585 gacagggcagccactlgtggcctggaaaaatcagatgtggaccaactttatctgataatcag 644
|||||
Db 301 GACAGGCCAGCCACTGTGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAG 360
QY 645 aacactag 652
|||||
Db 361 AACACTAG 368

RESULT 10
AC079615
LOCUS AC079615 80362 bp DNA PRI 03-JUL-2001
DEFINITION Homo sapiens clone Rp11-46216, complete sequence.
ACCESSION AC079615
VERSION AC079615.6 GI:14589743

KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 80362)
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 80362)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (04-SEP-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 3 (bases 1 to 80362)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (03-JUL-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Jul 3, 2001 this sequence version replaced gi:14018142.
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0462L06
----- Location/Qualifiers -----
source 1. .80362
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-462L6"
BASE COUNT 22997 a 15643 c 16118 g 25604 t
ORIGIN
Query Match 19.0%; Score 198; DB 9; Length 80362;
Best Local Similarity 100.0%; Pred. No. 1.8e-98;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 455 ttccaggagatgcctacaccgattatgtagctacgagatggtaccgagctcctgaacttc 514
|||||
Db 73475 TTCCAGGAGATGCCTACACCGATTATGTAGCTACGAGATGGTACCGAGCTCCTGAAC TTC 73534
QY 515 ttgtgggagatactcagtatggttcttcagtcgatatatgggctattggtgtgtttttg 574
|||||
Db 73535 TTGTGGGAGATACTCAGTATGGTTCCTTCAGTCGATATGGGTATTGGTGTGTTTGTG 73594
QY 575 cagagctcctgacaggccagccactgtggcctgggaaatcagatgtggaccaactttatc 634
|||||
Db 73595 CAGAGCTCCTGACAGGCCAGCCACTGTGGCCTGGAAATCAGATGTGGACCAACTTTATC 73654
QY 635 tgataatcagaacactag 652
|||||
Db 73655 TGATAATCAGAACACTAG 73672
RESULT 11
AC037489
ID AC037489 standard; DNA; HTG; 155028 BP.
XX
AC AC037489;
XX
SV AC037489.2
XX
DT 10-APR-2000 (Rel. 63, Created)
DT 14-MAY-2000 (Rel. 63, Last updated, Version 2)
XX
DE Homo sapiens chromosome 2 clone RP11-603F24 map 2, WORKING DRAFT SEQUENCE,
DE 26 unordered pieces.

XX HTG; HTGS_DRAFT; HTGS_PHASE1.
KW
XX Homo sapiens (human)
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX [1]
RN 1-155028
RP Birren B., Linton L., Nusbaum C., Lander E.;
RA "Homo sapiens chromosome 2, clone RP11-603F24";
RT Unpublished.
RL
XX [2]
RN 1-155028
RP Birren B., Linton L., Nusbaum C., Lander E., Abraham H., Allen N.,
RA Anderson S., Baldwin J., Barna N., Bastien V., Beda F., Boguslavskiy L.,
RA Boukhgalter B., Brown A., Burkett G., Campopiano A., Castle A., Choepel Y.,
RA Colangelo M., Collins S., Collymore A., Cooke P., DeArellano K., Dewar K.,
RA Diaz J.S., Dodge S., Domino M., Doyle M., Ferreira P., FitzHugh W.,
RA Gage D., Galagan J., Gardyna S., Ginde S., Goyette M., Graham L.,
RA Grand-Pierre N., Grant G., Hagos B., Heaford A., Horton L., Howland J.C.,
RA Iliev I., Johnson R., Jones C., Kann L., Karatas A., Klein J., LaRocque K.,
RA Lamazares R., Landers T., Lehoczyk J., Levine R., Lieu C., Liu G.,
RA Locke K., Macdonald P., Marquis N., McCarthy M., McEwan P., McGurk A.,
RA McKernan K., McPheeters R., Meldrim J., Meneus L., Mihova T., Miranda C.,
RA Mlenga V., Morrow J., Murphy T., Naylor J., Norman C.H., O'Connor T.,
RA O'Donnell P., O'Neill D., Oliver T.M., Oliver J., Peterson K., Pierre N.,
RA Pisani C., Pollara V., Raymond C., Riley R., Rogov P., Rothman D., Roy A.,
RA Santos R., Schauer S., Severy P., Spencer B., Stange-Thomann N.,
RA Stojanovic N., Subramanian A., Talamas J., Tesfaye S., Theodore J.,
RA Tirrell A., Travers M., Trigilio J., Vassiliev H., Viel R., Vo A.,
RA Wilson B., Wu X., Wyman D., Ye W.J., Young G., Zainoun J., Zimmer A.,
RA Zody M.;
RT
RL Submitted (09-APR-2000) to the EMBL/GenBank/DBJ databases.
RL Whitehead Institute/MIT Center for Genome Research, 320 Charles Street,
RL Cambridge, MA 02141, USA
XX
CC On May 12, 2000 this sequence version replaced gi:7528398.
CC All repeats were identified using RepeatMasker:
CC Smit, A.F.A. & Green, P. (1996-1997)
CC http://ftp.genome.washington.edu/RM/RepeatMasker.html
CC ----- Genome Center -----
CC Center: Whitehead Institute/ MIT Center for Genome Research
CC Center code: WIBR
CC Web site: http://www-seq.wi.mit.edu
CC Contact: sequence_submissions@genome.wi.mit.edu
CC ----- Project Information -----
CC Center project name: L9434
CC Center clone name: 603_F_24
CC ----- Summary Statistics -----
CC Sequencing vector: M13; M77815; 100% of reads
CC Chemistry: Dye-terminator Big Dye; 100% of reads
CC Assembly program: Phrap; version 0.960731
CC Consensus quality: 142960 bases at least Q40
CC Consensus quality: 148315 bases at least Q30
CC Consensus quality: 150779 bases at least Q20
CC Insert size: 164000; agarose-fp
CC Insert size: 152528; sum-of-contigs
CC Quality coverage: 4.4 in Q20 bases; agarose-fp
CC Quality coverage: 4.7 in Q20 bases; sum-of-contigs
CC -----
CC * NOTE: This is a 'working draft' sequence. It currently
CC * consists of 26 contigs. The true order of the pieces
CC * is not known and their order in this sequence record is
CC * arbitrary. Gaps between the contigs are represented as
CC * runs of N, but the exact sizes of the gaps are unknown.
CC * This record will be updated with the finished sequence
CC * as soon as it is available and the accession number will
CC * be preserved.
CC 1 1104: contig of 1104 bp in length
CC * 1105 1204: gap of 100 bp
CC *

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 164281)
Waterston,R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 164281)
Waterston,R.H.
Direct Submission
Submitted (19-JUL-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Jul 19, 2001 this sequence version replaced gi:7770604.

COMMENT

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
----- Project Information -----
Center project name: H_NH0603F24
Drafting center: WIBR

----- Summary Statistics -----
Sequencing vector: M13; 47%
Sequencing vector: plasmid; 52%
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 162138 bases at least Q40
Consensus quality: 163048 bases at least Q30
Consensus quality: 163441 bases at least Q20
Insert size: 169000; agarose-fp
Insert size: 165355; sum-of-contigs
Quality coverage: 8.62 in Q20 bases; agarose-fp
Quality coverage: 8.84 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 16967: contig of 16967 bp in length
* 16968 17067: gap of unknown length
* 17068 48462: contig of 31395 bp in length
* 48463 48562: gap of unknown length
* 48563 99374: contig of 50812 bp in length
* 99375 99474: gap of unknown length
* 99475 164281: contig of 64807 bp in length.

FEATURES

source
1. .164281
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/clone="RP11-603F24"
1. .16967
/note="assembly_name:Contig8"
17068. .48462
/note="assembly_name:Contig9"
48563. .99374
/note="assembly_name:Contig10
clone_end:T7
vector_side:right"
99475. .164281
/note="assembly_name:Contig11
clone_end:SP6
vector_side:left"
46622 a 33513 c 33986 g 49860 t 300 others

BASE COUNT 46622 a 33513 c 33986 g 49860 t 300 others
ORIGIN

Query Match 19.0%; Score 198; DB 2; Length 164281;
Best Local Similarity 100.0%; Pred. No. 1.8e-98;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 455 ttccaggagatgcctacaccgattatgttagctagagatggtaccgagctcctgaaacttc 514
Db 142462 TTCCAGGAGATGCCCTACACCGATTATGTAGCTACGAGATGGTACCGAGCTCCTGAAC TTC 142521
QY 515 ttgtgggagatactcagtagtgggtttcagtcagtcagatatatgggctattggtgtgttttg 574
Db 142522 TTGTGGGAGATACTCAGTATGGTCTTCAGTCGATATATGGCTATTGGTGTGTTTTTG 142581
QY 575 cagagctcctgcagggccagccactgtggcctggaaaatcagatgtggaccaactttatc 634
Db 142582 CAGAGCTCCTGACAGGCCAGCCACTGTGGCCTGGAAAATCAGATGTGGACCACTTTATC 142641
QY 635 tgataatcagaacactag 652
Db 142642 TGATAATCAGAACACTAG 142659

RESULT 13

AC037489/C standard; DNA; HTG; 155028 BP.

AC037489;

AC037489.2

10-APR-2000 (Rel. 63, Created)
14-MAY-2000 (Rel. 63, Last updated, Version 2)

Homo sapiens chromosome 2 clone RP11-603F24 map 2, WORKING DRAFT SEQUENCE,
26 unordered pieces.

HTG; HTGS_DRAFT; HTGS_PHASE1.

Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.

[1]
1-155028
Birren B., Linton L., Nusbaum C., Lander E.;

"Homo sapiens chromosome 2, clone RP11-603F24";
Unpublished.

[2]

Birren B., Linton L., Nusbaum C., Lander E., Abraham H., Allen N.,
Anderson S., Baldwin J., Barna N., Bastien V., Bida F., Boguslavsky L.,
Boukhgalter B., Brown A., Burkett G., Campopiano A., Castle A., Choepel Y.,
Colangelo M., Collins S., Collymore A., Cooke P., Dearellano K., Dewar K.,
Diaz J.S., Dodge S., Domino M., Doyle M., Ferreira P., FitzHugh W.,
Gage D., Galagan J., Gardyna S., Ginde S., Goyette M., Graham L.,
Grand-pierre N., Grant G., Hagos B., Headford A., Horton L., Howland J.C.,
Illiev I., Johnson R., Jones C., Kann L., Karatas A., Klein J., LaRocque K.,
Lamazares R., Landers T., Lehoczy J., Levine R., Lieu C., Liu G.,
Locke K., Macdonald P., Marquis N., McCarthy M., McEwan P., McGurk A.,
McKernan K., McPheeters R., Meldrim J., Meneus L., Mihova T., Miranda C.,
Mlenga V., Morrow J., Murphy T., Naylor J., Norman C.H., O'Connor T.,
O'Donnell P., O'Neil D., Oliver T.M., Oliver J., Peterson K., Pierre N.,
Pisani C., Pollara V., Raymond C., Riley R., Rogov P., Rothman D., Roy A.,
Santos R., Schauer S., Severy P., Spencer B., Stange-Thomann N.,
Stojanovic N., Subramanian A., Talamas J., Tesfaye S., Theodore J.,
Tirrell A., Travers M., Trigilio J., Vassiliev H., Viel R., Vo A.,
Wilson B., Wu X., Wyman D., Ye W.J., Young G., Zainoun J., Zimmer A.,
Zody M.;

; Submitted (09-APR-2000) to the EMBL/GenBank/DBJ databases.
Whitehead Institute/MIT Center for Genome Research, 320 Charles Street,
Cambridge, MA 02141, USA

```
CC On May 12, 2000 this sequence version replaced gi:7528398.
CC All repeats were identified using RepeatMasker:
CC Smit, A.F.A. & Green, P. (1996-1997)
CC http://ftp.genome.washington.edu/RM/RepeatMasker.html
CC ----- Genome Center
CC Center: Whitehead Institute/ MIT Center for Genome Research
CC Center code: WTHR
CC Web site: http://www-seq.wi.mit.edu
CC Contact: sequence_submissions@genome.wi.mit.edu
CC ----- Project Information
CC Center project name: L9434
CC Center clone name: 603_F_24
CC ----- Summary Statistics
CC Sequencing vector: M13; M77815; 100% of reads
CC Chemistry: Dye-terminator Big Dye; 100% of reads
CC Assembly program: Phrap; version 0.960731
CC Consensus quality: 142960 bases at least Q40
CC Consensus quality: 148315 bases at least Q30
CC Consensus quality: 150779 bases at least Q20
CC Insert size: 164000; agarose-fp
CC Insert size: 152528; sum-of-contigs
CC Quality coverage: 4.4 in Q20 bases; agarose-fp
CC Quality coverage: 4.7 in Q20 bases; sum-of-contigs
CC -----
CC * NOTE: This is a 'working draft' sequence. It currently
CC * consists of 26 contigs. The true order of the pieces
CC * is not known and their order in this sequence record is
CC * arbitrary. Gaps between the contigs are represented as
CC * runs of N, but the exact sizes of the gaps are unknown.
CC * This record will be updated with the finished sequence
CC * as soon as it is available and the accession number will
CC * be preserved.
CC *
CC * 1 1104: contig of 1104 bp in length
CC * 1105 1204: gap of 100 bp
CC * 1205 2330: contig of 1126 bp in length
CC * 2331 2430: gap of 100 bp
CC * 2431 4965: contig of 2535 bp in length
CC * 4966 5065: gap of 100 bp
CC * 5066 7346: contig of 2281 bp in length
CC * 7347 7446: gap of 100 bp
CC * 7447 9679: contig of 2233 bp in length
CC * 9680 9779: gap of 100 bp
CC * 9780 11977: contig of 2198 bp in length
CC * 11978 12077: gap of 100 bp
CC * 12078 14038: contig of 1961 bp in length
CC * 14039 14138: gap of 100 bp
CC * 14139 15917: contig of 1779 bp in length
CC * 15918 16017: gap of 100 bp
CC * 16018 18776: contig of 2759 bp in length
CC * 18777 18876: gap of 100 bp
CC * 18877 21090: contig of 2214 bp in length
CC * 21091 21190: gap of 100 bp
CC * 21191 23923: contig of 2733 bp in length
CC * 23924 24023: gap of 100 bp
CC * 24024 27664: contig of 3641 bp in length
CC * 27665 27764: gap of 100 bp
CC * 27765 31535: contig of 3771 bp in length
CC * 31536 31635: gap of 100 bp
CC * 31636 36572: contig of 4937 bp in length
CC * 36573 36672: gap of 100 bp
CC * 36673 43155: contig of 6483 bp in length
CC * 43156 43255: gap of 100 bp
CC * 43256 48604: contig of 5349 bp in length
CC * 48605 48704: gap of 100 bp
CC * 48705 54578: contig of 5874 bp in length
CC * 54579 54678: gap of 100 bp
CC * 54679 61303: contig of 6625 bp in length
CC * 61304 61403: gap of 100 bp
CC * 61404 67988: contig of 6585 bp in length
CC * 67989 68088: gap of 100 bp
CC * 68089 75061: contig of 6973 bp in length
CC * 75062 75161: gap of 100 bp
CC * 75162 84322: contig of 9161 bp in length
CC *
```

```
CC * 84323 84422: gap of 100 bp
CC * 84423 93345: contig of 8923 bp in length
CC * 93346 93445: gap of 100 bp
CC * 93446 107415: contig of 13970 bp in length
CC * 107416 107515: gap of 100 bp
CC * 107516 119348: contig of 11833 bp in length
CC * 119349 119448: gap of 100 bp
CC * 119449 134591: contig of 15143 bp in length
CC * 134592 134691: gap of 100 bp
CC * 134692 155028: contig of 20337 bp in length.
CC XX
FH Key Location/Qualifiers
FH source 1. .155028
FT /chromosome="2"
FT /db_xref="taxon:9606"
FT /organism="Homo sapiens"
FT /map="2"
FT /clone="RP11-603F24"
FT /clone_lib="RPC1-11 Human Male BAC"
FT 1. .1104
FT /note="assembly_fragment"
FT 1205. .2330
FT /note="assembly_fragment"
FT 2431. .4965
FT /note="assembly_fragment"
FT 5066. .7346
FT /note="assembly_fragment"
FT 7447. .9679
FT /note="assembly_fragment"
FT 9780. .11977
FT /note="assembly_fragment"
FT 12078. .14038
FT /note="assembly_fragment"
FT 14139. .15917
FT /note="assembly_fragment"
FT 16018. .18776
FT /note="assembly_fragment"
FT 18877. .21090
FT /note="assembly_fragment"
FT 21191. .23923
FT /note="assembly_fragment"
FT 24024. .27664
FT /note="assembly_fragment"
FT 27765. .31535
FT /note="assembly_fragment"
FT 31636. .36572
FT /note="assembly_fragment"
FT 36673. .43155
FT /note="assembly_fragment"
FT 43256. .48604
FT /note="assembly_fragment"
FT 48705. .54578
FT /note="assembly_fragment"
FT 54679. .61303
FT /note="assembly_fragment"
FT 61404. .67988
FT /note="assembly_fragment"
FT 68089. .75061
FT /note="assembly_fragment"
FT 75162. .84322
FT /note="assembly_fragment"
FT 84423. .93345
FT /note="assembly_fragment clone_end:T7 vector_side:right"
FT 93446. .107415
FT /note="assembly_fragment"
FT 107516. .119348
FT /note="assembly_fragment"
FT 119449. .134591
FT /note="assembly_fragment clone_end:SP6 vector_side:left"
FT 134692. .155028
FT /note="assembly_fragment"
FT XX
```


Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacggggttcgccatgttgaccaggctggtctcgaaactc 696
|||||
Db 9035 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9081

RESULT 17
GGO312528 10135 bp DNA PRI 13-FEB-2001
LOCUS Gorilla gorilla gorilla Xql3.3 chromosome non-coding sequence,
DEFINITION isolate G1979.
ACCESSION AJ312528
VERSION AJ312528.1 GI:12830524
KEYWORDS
SOURCE lowland gorilla.
ORGANISM Gorilla gorilla gorilla
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
REFERENCE 1 (bases 1 to 10135)
AUTHORS Kaessmann,H., Wiebe,V., Weiss,G. and Paabo,S.
TITLE Great ape DNA sequences reveal a reduced diversity and an expansion
in humans
JOURNAL Nat. Genet. 27 (2), 155-156 (2001)
PUBMED 11175781
AUTHORS Kaessmann,H.G.F.
TITLE Direct Submission
JOURNAL Submitted (09-FEB-2001) Kaessmann H.G.F., Evolutionary Genetics
(Prof. Paabo), Max-Planck-Institute for Evolutionary Anthropology,
Inselstrasse 22, Leipzig, Saxony, 04103, GERMANY
COMMENT non-coding sequence.
FEATURES Location/Qualifiers
source 1..10135
/organism="Gorilla gorilla gorilla"
/isolate="G1979"
/sub_species="gorilla"
/db_xref="taxon:9595"
/chromosome="X"
/map="Xql3.3"
/sex="male"

BASE COUNT 3233 a 2045 c 1802 g 3055 t
ORIGIN

Query Match 4.5%; Score 47; DB 9; Length 10135;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacggggttcgccatgttgaccaggctggtctcgaaactc 696
|||||
Db 9035 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9081

RESULT 18
GGO241093 10139 bp DNA PRI 07-MAY-1999
LOCUS Gorilla gorilla Xql3.3 non-coding region, isolated from Western
DEFINITION lowland.
ACCESSION AJ241093
VERSION AJ241093.1 GI:4775320
KEYWORDS non-coding region.
SOURCE gorilla.
ORGANISM Gorilla gorilla
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
REFERENCE 1 (bases 1 to 10139)
AUTHORS Kaessmann,H., Heissig,F., von Haeseler,A. and Paabo,S.
TITLE DNA sequence variation in a non-coding region of low recombination
on the human X chromosome
JOURNAL Nat. Genet. 22 (1), 78-81 (1999)
MEDLINE 99251584
REFERENCE 2 (bases 1 to 10139)

AUTHORS Kaessmann,H.
TITLE Direct Submission
JOURNAL Submitted (22-MAR-1999) Kaessmann H, Max-Planck-Institute of
Evolutionary Anthropology Inselstrasse 22, 04103 Leipzig, FRG
FEATURES Location/Qualifiers
source 1..10139
/organism="Gorilla gorilla"
/isolate="from Western lowland"
/db_xref="taxon:9593"
/chromosome="X"
/map="Xql3.3"
misc_feature 1..10139
/note="non-coding region"

BASE COUNT 3234 a 2044 c 1803 g 3058 t
ORIGIN

Query Match 4.5%; Score 47; DB 9; Length 10139;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacggggttcgccatgttgaccaggctggtctcgaaactc 696
|||||
Db 9038 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9084

RESULT 19
GGO312522 10139 bp DNA PRI 13-FEB-2001
LOCUS Gorilla gorilla gorilla Xql3.3 chromosome non-coding sequence,
DEFINITION isolate G107W.
ACCESSION AJ312522
VERSION AJ312522.1 GI:12830518
KEYWORDS
SOURCE lowland gorilla.
ORGANISM Gorilla gorilla gorilla
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
REFERENCE 1 (bases 1 to 10139)
AUTHORS Kaessmann,H., Wiebe,V., Weiss,G. and Paabo,S.
TITLE Great ape DNA sequences reveal a reduced diversity and an expansion
in humans
JOURNAL Nat. Genet. 27 (2), 155-156 (2001)
PUBMED 11175781
REFERENCE 2 (bases 1 to 10139)
AUTHORS Kaessmann,H.G.F.
TITLE Direct Submission
JOURNAL Submitted (09-FEB-2001) Kaessmann H.G.F., Evolutionary Genetics
(Prof. Paabo), Max-Planck-Institute for Evolutionary Anthropology,
Inselstrasse 22, Leipzig, Saxony, 04103, GERMANY
COMMENT non-coding sequence.
FEATURES Location/Qualifiers
source 1..10139
/organism="Gorilla gorilla gorilla"
/isolate="G107W"
/sub_species="gorilla"
/db_xref="taxon:9595"
/chromosome="X"
/map="Xql3.3"
/sex="male"

BASE COUNT 3238 a 2043 c 1797 g 3061 t
ORIGIN

Query Match 4.5%; Score 47; DB 9; Length 10139;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacggggttcgccatgttgaccaggctggtctcgaaactc 696
|||||
Db 9039 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9085

```
RESULT 20
GGO312529      10139 bp      DNA      PRI      13-FEB-2001
LOCUS
DEFINITION
Gorilla gorilla gorilla Xql3.3 chromosome non-coding sequence,
isolate G014.
ACCESSION
AJ312529
VERSION
AJ312529.1 GI:12830525
KEYWORDS
SOURCE
lowland gorilla.
ORGANISM
Gorilla gorilla gorilla
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
REFERENCE
1 (bases 1 to 10139)
AUTHORS
Kaessmann,H., Wiebe,V., Weiss,G. and Paabo,S.
TITLE
Great ape DNA sequences reveal a reduced diversity and an expansion
in humans
JOURNAL
PUBMED
Nat. Genet. 27 (2), 155-156 (2001)
REFERENCE
2 (bases 1 to 10139)
AUTHORS
Kaessmann,H.G.F.
TITLE
Direct Submission
JOURNAL
Submitted (09-FEB-2001) Kaessmann H.G.F., Evolutionary Genetics
(Prof. Paabo), Max-Planck-Institute for Evolutionary Anthropology,
Inselstrasse 22, Leipzig, Saxony, 04103, GERMANY
COMMENT
non-coding sequence.
FEATURES
Location/Qualifiers
source
1..10139
/organism="Gorilla gorilla gorilla"
/isolate="G014"
/sub_species="gorilla"
/db_xref="taxon:9595"
/chromosome="X"
/map="Xql3.3"
/sex="male"
BASE COUNT      3238 a      2043 c      1797 g      3061 t
ORIGIN

Query Match      4.5%; Score 47; DB 9; Length 10139;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacggggttcgccatgttgaccaggctggtcgaactc 696
|||||
Db 9039 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9085

RESULT 21
GGO312530      10139 bp      DNA      PRI      13-FEB-2001
LOCUS
DEFINITION
Gorilla gorilla gorilla Xql3.3 chromosome non-coding sequence,
isolate G1530.
ACCESSION
AJ312530
VERSION
AJ312530.1 GI:12830526
KEYWORDS
SOURCE
lowland gorilla.
ORGANISM
Gorilla gorilla gorilla
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
REFERENCE
1 (bases 1 to 10139)
AUTHORS
Kaessmann,H., Wiebe,V., Weiss,G. and Paabo,S.
TITLE
Great ape DNA sequences reveal a reduced diversity and an expansion
in humans
JOURNAL
PUBMED
Nat. Genet. 27 (2), 155-156 (2001)
REFERENCE
2 (bases 1 to 10139)
AUTHORS
Kaessmann,H.G.F.
TITLE
Direct Submission
JOURNAL
Submitted (09-FEB-2001) Kaessmann H.G.F., Evolutionary Genetics
(Prof. Paabo), Max-Planck-Institute for Evolutionary Anthropology,
Inselstrasse 22, Leipzig, Saxony, 04103, GERMANY
COMMENT
non-coding sequence.
FEATURES
Location/Qualifiers
source
1..10139
/organism="Gorilla gorilla gorilla"
/isolate="G014"
/sub_species="gorilla"
/db_xref="taxon:9595"
/chromosome="X"
/map="Xql3.3"
/sex="male"
BASE COUNT      3238 a      2043 c      1797 g      3061 t
ORIGIN

Query Match      4.5%; Score 47; DB 9; Length 10139;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacggggttcgccatgttgaccaggctggtcgaactc 696
|||||
Db 9039 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9085

RESULT 23
GGO312523      10142 bp      DNA      PRI      13-FEB-2001
LOCUS
DEFINITION
Gorilla gorilla gorilla Xql3.3 chromosome non-coding sequence,
isolate G167W.
COMMENT
non-coding sequence.
FEATURES
Location/Qualifiers
source
1..10141
/organism="Gorilla gorilla gorilla"
/isolate="G1537W"
/sub_species="gorilla"
/db_xref="taxon:9595"
/chromosome="X"
/map="Xql3.3"
/sex="male"
BASE COUNT      3234 a      2048 c      1802 g      3057 t
ORIGIN

Query Match      4.5%; Score 47; DB 9; Length 10141;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacggggttcgccatgttgaccaggctggtcgaactc 696
|||||
Db 9041 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9087

RESULT 23
GGO312523      10142 bp      DNA      PRI      13-FEB-2001
LOCUS
DEFINITION
Gorilla gorilla gorilla Xql3.3 chromosome non-coding sequence,
isolate G167W.
```

```
source
1..10139
/organism="Gorilla gorilla gorilla"
/isolate="G1530"
/sub_species="gorilla"
/db_xref="taxon:9595"
/chromosome="X"
/map="Xql3.3"
/sex="male"
BASE COUNT      3238 a      2043 c      1797 g      3061 t
ORIGIN

Query Match      4.5%; Score 47; DB 9; Length 10139;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacggggttcgccatgttgaccaggctggtcgaactc 696
|||||
Db 9039 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9085

RESULT 22
GGO312521      10141 bp      DNA      PRI      13-FEB-2001
LOCUS
DEFINITION
Gorilla gorilla gorilla Xql3.3 chromosome non-coding sequence,
isolate G1537W.
ACCESSION
AJ312521
VERSION
AJ312521.1 GI:12830517
KEYWORDS
SOURCE
lowland gorilla.
ORGANISM
Gorilla gorilla gorilla
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
REFERENCE
1 (bases 1 to 10141)
AUTHORS
Kaessmann,H., Wiebe,V., Weiss,G. and Paabo,S.
TITLE
Great ape DNA sequences reveal a reduced diversity and an expansion
in humans
JOURNAL
PUBMED
Nat. Genet. 27 (2), 155-156 (2001)
REFERENCE
2 (bases 1 to 10141)
AUTHORS
Kaessmann,H.G.F.
TITLE
Direct Submission
JOURNAL
Submitted (09-FEB-2001) Kaessmann H.G.F., Evolutionary Genetics
(Prof. Paabo), Max-Planck-Institute for Evolutionary Anthropology,
Inselstrasse 22, Leipzig, Saxony, 04103, GERMANY
COMMENT
non-coding sequence.
FEATURES
Location/Qualifiers
source
1..10141
/organism="Gorilla gorilla gorilla"
/isolate="G1537W"
/sub_species="gorilla"
/db_xref="taxon:9595"
/chromosome="X"
/map="Xql3.3"
/sex="male"
BASE COUNT      3234 a      2048 c      1802 g      3057 t
ORIGIN

Query Match      4.5%; Score 47; DB 9; Length 10141;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacggggttcgccatgttgaccaggctggtcgaactc 696
|||||
Db 9041 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9087

RESULT 23
GGO312523      10142 bp      DNA      PRI      13-FEB-2001
LOCUS
DEFINITION
Gorilla gorilla gorilla Xql3.3 chromosome non-coding sequence,
isolate G167W.
```

ACCESSION AJ312523
VERSION AJ312523.1 GI:12830519
KEYWORDS
SOURCE
ORGANISM lowland gorilla.
Gorilla gorilla gorilla
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
REFERENCE 1 (bases 1 to 10142)
AUTHORS Kaessmann,H., Wiebe,V., Weiss,G. and Paabo,S.
TITLE Great ape DNA sequences reveal a reduced diversity and an expansion in humans
JOURNAL Nat. Genet. 27 (2), 155-156 (2001)
PUBMED 11175781
REFERENCE 2 (bases 1 to 10142)
AUTHORS Kaessmann,H.G.F.
TITLE Direct Submission
JOURNAL Submitted (09-FEB-2001) Kaessmann H.G.F., Evolutionary Genetics (Prof. Paabo), Max-Planck-Institute for Evolutionary Anthropology, Inselstrasse 22, Leipzig, Saxony, 04103, GERMANY
COMMENT non-coding sequence.
FEATURES
source Location/Qualifiers
1..10142
/organism="Gorilla gorilla gorilla"
/isolate="G167W"
/sub_species="gorilla"
/db_xref="taxon:9595"
/chromosome="X"
/map="Xq13.3"
/sex="male"
BASE COUNT 3237 a 2044 c 1799 g 3062 t
ORIGIN
Query Match 4.5%; Score 47; DB 9; Length 10142;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 650 tagtagagacgggttcgccatgttgaccaggctggtctcgaactc 696
|||||
Db 9042 TAGTAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9088
RESULT 24
GGO312525
LOCUS
DEFINITION Gorilla gorilla beringei Xq13.3 chromosome non-coding sequence, isolate Mou074. PRI 13-FEB-2001
ACCESSION AJ312525
VERSION AJ312525.1 GI:12830521
KEYWORDS
SOURCE
ORGANISM mountain gorilla.
Gorilla gorilla beringei
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
REFERENCE 1 (bases 1 to 10142)
AUTHORS Kaessmann,H., Wiebe,V., Weiss,G. and Paabo,S.
TITLE Great ape DNA sequences reveal a reduced diversity and an expansion in humans
JOURNAL Nat. Genet. 27 (2), 155-156 (2001)
PUBMED 11175781
REFERENCE 2 (bases 1 to 10142)
AUTHORS Kaessmann,H.G.F.
TITLE Direct Submission
JOURNAL Submitted (09-FEB-2001) Kaessmann H.G.F., Evolutionary Genetics (Prof. Paabo), Max-Planck-Institute for Evolutionary Anthropology, Inselstrasse 22, Leipzig, Saxony, 04103, GERMANY
COMMENT non-coding sequence.
FEATURES
source Location/Qualifiers
1..10142
/organism="Gorilla gorilla beringei"
/isolate="Mou074"
/sub_species="beringei"
/db_xref="taxon:9594"

/chromosome="X"
/map="Xq13.3"
/sex="male"
BASE COUNT 3238 a 2044 c 1799 g 3061 t
ORIGIN
Query Match 4.5%; Score 47; DB 9; Length 10142;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 650 tagtagagacgggttcgccatgttgaccaggctggtctcgaactc 696
|||||
Db 9042 TAGTAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9088
RESULT 25
PPA270090
LOCUS
DEFINITION Pan paniscus Xq13.3 non-coding region, B1. PRI 05-NOV-1999
ACCESSION AJ270090
VERSION AJ270090.1 GI:6273615
KEYWORDS non-coding region.
SOURCE pygmy chimpanzee.
ORGANISM Pan paniscus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 10145)
AUTHORS Kaessmann,H., Wiebe,V. and Paabo,S.
TITLE Extensive nuclear DNA sequence diversity among chimpanzees
JOURNAL Science 286 (5442), 1159-1162 (1999)
MEDLINE 20018332
REFERENCE 2 (bases 1 to 10145)
AUTHORS Kaessmann,H.G.F.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo, Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22, Leipzig, 04103, GERMANY
FEATURES
source Location/Qualifiers
1..10145
/organism="Pan paniscus"
/db_xref="taxon:9597"
/country="Democratic Republic of the Congo"
/chromosome="X"
/sex="male"
/map="Xq13.3"
/note="B1"
misc_feature 1..10145
/note="non-coding DNA"
BASE COUNT 3226 a 2031 c 1812 g 3076 t
ORIGIN
Query Match 4.5%; Score 47; DB 9; Length 10145;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 650 tagtagagacgggttcgccatgttgaccaggctggtctcgaactc 696
|||||
Db 9045 TAGTAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9091
RESULT 26
PPA270091
LOCUS
DEFINITION Pan paniscus Xq13.3 non-coding region, B4. PRI 05-NOV-1999
ACCESSION AJ270091
VERSION AJ270091.1 GI:6273616
KEYWORDS non-coding region.
SOURCE pygmy chimpanzee.
ORGANISM Pan paniscus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

us-09-671-050-9.oli20.rge

Fri May 3 11:02:09 2002

Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

REFERENCE 1 (bases 1 to 10145)
AUTHORS Kaessmann,H., Wiebe,V. and Paabo,S.
TITLE Extensive nuclear DNA sequence diversity among chimpanzees
JOURNAL Science 286 (5442), 1159-1162 (1999)
MEDLINE 20018332
REFERENCE 2 (bases 1 to 10145)
AUTHORS Kaessmann,H.G.F.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY

FEATURES
source
1..10145
/organism="Pan paniscus"
/db_xref="taxon:9597"
/country="Democratic Republic of the Congo"
/chromosome="X"
/sex="male"
/map="Xq13.3"
/note="B4"

misc_feature 1..10145
/note="non-coding DNA"
BASE COUNT 3224 a 2032 c 1812 g 3077 t
ORIGIN

Query Match 4.5%; Score 47; DB 9; Length 10145;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacggggttcgccatgttgaccaggctggtctcgaactc 696
|||||
Db 9045 TAGTAGAGACGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9091

RESULT 27
PPA270093 10145 bp DNA PRI 05-NOV-1999
LOCUS Pan paniscus Xq13.3 non-coding region, B2.
DEFINITION
ACCESSION AJ270093
VERSION AJ270093.1 GI:6273618
KEYWORDS non-coding region.
SOURCE pygmy chimpanzee.

ORGANISM Pan paniscus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 10145)
AUTHORS Kaessmann,H., Wiebe,V. and Paabo,S.
TITLE Extensive nuclear DNA sequence diversity among chimpanzees
JOURNAL Science 286 (5442), 1159-1162 (1999)
MEDLINE 20018332

REFERENCE 2 (bases 1 to 10145)
AUTHORS Kaessmann,H.G.F.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY

FEATURES
source
1..10145
/organism="Pan paniscus"
/db_xref="taxon:9597"
/country="Democratic Republic of the Congo"
/chromosome="X"
/sex="male"
/map="Xq13.3"
/note="B2"

misc_feature 1..10145
/note="non-coding DNA"
BASE COUNT 3224 a 2032 c 1812 g 3077 t
ORIGIN

Query Match 4.5%; Score 47; DB 9; Length 10145;

QY 650 tagtagagacggggttcgccatgttgaccaggctggtctcgaactc 696
|||||
Db 9045 TAGTAGAGACGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9091

RESULT 28
PPA270094 10145 bp DNA PRI 05-NOV-1999
LOCUS Pan paniscus Xq13.3 non-coding region, B3.
DEFINITION
ACCESSION AJ270094
VERSION AJ270094.1 GI:6273638
KEYWORDS non-coding region.
SOURCE pygmy chimpanzee.

ORGANISM Pan paniscus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

REFERENCE 1 (bases 1 to 10145)
AUTHORS Kaessmann,H., Wiebe,V. and Paabo,S.
TITLE Extensive nuclear DNA sequence diversity among chimpanzees
JOURNAL Science 286 (5442), 1159-1162 (1999)
MEDLINE 20018332
REFERENCE 2 (bases 1 to 10145)
AUTHORS Kaessmann,H.G.F.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY

FEATURES
source
1..10145
/organism="Pan paniscus"
/db_xref="taxon:9597"
/country="Democratic Republic of the Congo"
/chromosome="X"
/sex="male"
/map="Xq13.3"
/note="B3"

misc_feature 1..10145
/note="non-coding DNA"
BASE COUNT 3224 a 2032 c 1812 g 3077 t
ORIGIN

Query Match 4.5%; Score 47; DB 9; Length 10145;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacggggttcgccatgttgaccaggctggtctcgaactc 696
|||||
Db 9045 TAGTAGAGACGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9091

RESULT 29
PPA270092 10151 bp DNA PRI 05-NOV-1999
LOCUS Pan paniscus Xq13.3 non-coding region, B5.
DEFINITION
ACCESSION AJ270092
VERSION AJ270092.1 GI:6273617
KEYWORDS non-coding region.
SOURCE pygmy chimpanzee.

ORGANISM Pan paniscus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

REFERENCE 1 (bases 1 to 10151)
AUTHORS Kaessmann,H., Wiebe,V. and Paabo,S.
TITLE Extensive nuclear DNA sequence diversity among chimpanzees
JOURNAL Science 286 (5442), 1159-1162 (1999)
MEDLINE 20018332
REFERENCE 2 (bases 1 to 10151)
AUTHORS Kaessmann,H.G.F.
TITLE Direct Submission

JOURNAL Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY

FEATURES source
Location/Qualifiers
1. .10151
/organism="Pan paniscus"
/db_xref="taxon:9597"
/country="Democratic Republic of the Congo"
/chromosome="X"
/sex="male"
/map="Xq13.3"
/note="B5"
1. .10151
/note="non-coding DNA"

misc_feature 3227 a 2032 c 1811 g 3081 t
BASE COUNT 3227 a 2032 c 1811 g 3081 t
ORIGIN

Query Match 4.5%; Score 47; DB 9; Length 10151;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacggggttcgccatgttgaccaggctggtctcgaaactc 696
|||||
Db 9051 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9097

RESULT 30
PTR270065
LOCUS PTR270065 10151 bp DNA PRI 05-NOV-1999
DEFINITION Pan troglodytes troglodytes Xq13.3 non-coding region, C2.
ACCESSION AJ270065
VERSION AJ270065.1 GI:6273630
KEYWORDS non-coding region.
SOURCE chimpanzee.
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
AUTHORS 1 (bases 1 to 10151)
TITLE Kaessmann,H., Wiebe,V. and Paabo,S.
JOURNAL Extensive nuclear DNA sequence diversity among chimpanzees
MEDLINE Science 286 (5442), 1159-1162 (1999)
REFERENCE 20018332
AUTHORS 2 (bases 1 to 10151)
TITLE Kaessmann,H.G.F.
JOURNAL Direct Submission
Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY
FEATURES source
Location/Qualifiers
1. .10151
/organism="Pan troglodytes"
/sub_species="troglodytes"
/db_xref="taxon:9598"
/country="Gabon"
/chromosome="X"
/sex="male"
/map="Xq13.3"
/note="C2"
1. .10151
/note="non-coding DNA"

misc_feature 3221 a 2033 c 1814 g 3083 t
BASE COUNT 3221 a 2033 c 1814 g 3083 t
ORIGIN

Query Match 4.5%; Score 47; DB 9; Length 10151;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacggggttcgccatgttgaccaggctggtctcgaaactc 696
|||||
Db 9051 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9097

RESULT 31
PTR270066
LOCUS PTR270066 10151 bp DNA PRI 05-NOV-1999
DEFINITION Pan troglodytes troglodytes Xq13.3 non-coding region, C3.
ACCESSION AJ270066
VERSION AJ270066.1 GI:6273631
KEYWORDS non-coding region.
SOURCE chimpanzee.
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
AUTHORS 1 (bases 1 to 10151)
TITLE Kaessmann,H., Wiebe,V. and Paabo,S.
JOURNAL Extensive nuclear DNA sequence diversity among chimpanzees
MEDLINE Science 286 (5442), 1159-1162 (1999)
REFERENCE 20018332
AUTHORS 2 (bases 1 to 10151)
TITLE Kaessmann,H.G.F.
JOURNAL Direct Submission
Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY
FEATURES source
Location/Qualifiers
1. .10151
/organism="Pan troglodytes"
/sub_species="troglodytes"
/db_xref="taxon:9598"
/country="Gabon"
/chromosome="X"
/sex="male"
/map="Xq13.3"
/note="C3"
1. .10151
/note="non-coding DNA"

misc_feature 3221 a 2031 c 1817 g 3082 t
BASE COUNT 3221 a 2031 c 1817 g 3082 t
ORIGIN

Query Match 4.5%; Score 47; DB 9; Length 10151;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacggggttcgccatgttgaccaggctggtctcgaaactc 696
|||||
Db 9051 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9097

RESULT 32
PTR270068
LOCUS PTR270068 10151 bp DNA PRI 05-NOV-1999
DEFINITION Pan troglodytes troglodytes Xq13.3 non-coding region, C7.
ACCESSION AJ270068
VERSION AJ270068.1 GI:6273633
KEYWORDS non-coding region.
SOURCE chimpanzee.
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
AUTHORS 1 (bases 1 to 10151)
TITLE Kaessmann,H., Wiebe,V. and Paabo,S.
JOURNAL Extensive nuclear DNA sequence diversity among chimpanzees
MEDLINE Science 286 (5442), 1159-1162 (1999)
REFERENCE 20018332
AUTHORS 2 (bases 1 to 10151)
TITLE Kaessmann,H.G.F.
JOURNAL Direct Submission
Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY
FEATURES source
Location/Qualifiers
1. .10151
/organism="Pan troglodytes"

misc_feature 1. .10151 /sub_species="troglodytes" /db_xref="taxon:9598" /country="Gabon" /chromosome="X" /sex="male" /map="Xql3.3" /note="C7" /note="non-coding DNA" 3225 a 2031 c 1813 g 3082 t

BASE COUNT 3225 a 2031 c 1813 g 3082 t

ORIGIN

Query Match 4.5%; Score 47; DB 9; Length 10151; Best Local Similarity 100.0%; Pred. No. 1.5e-14; Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 650 tagtagagacgggttcgccatgttgaccaggctggtctcgaaactc 696
|||||

Db 9051 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9097

RESULT 33

PTR270070 10151 bp DNA PRI 05-NOV-1999

LOCUS Pan troglodytes troglodytes Xql3.3 non-coding region, C4.

DEFINITION AJ270070

ACCESSION AJ270070

VERSION AJ270070.1 GI:6273635

KEYWORDS non-coding region.

SOURCE chimpanzee.

ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 (bases 1 to 10151)
Kaessmann,H., Wiebe,V. and Paabo,S.
Extensive nuclear DNA sequence diversity among chimpanzees
Science 286 (5442), 1159-1162 (1999)

REFERENCE 2 (bases 1 to 10151)
Kaessmann,H.G.F.
Direct Submission
Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY

FEATURES
source
1. .10151
/organism="Pan troglodytes"
/sub_species="troglodytes"
/db_xref="taxon:9598"
/country="Gabon"
/chromosome="X"
/sex="male"
/map="Xql3.3"
/note="C8"
/note="non-coding DNA"
1. .10151
/note="non-coding DNA" 3221 a 2032 c 1818 g 3080 t

BASE COUNT 3221 a 2032 c 1818 g 3080 t

ORIGIN

Query Match 4.5%; Score 47; DB 9; Length 10151; Best Local Similarity 100.0%; Pred. No. 1.5e-14; Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacgggttcgccatgttgaccaggctggtctcgaaactc 696
|||||

Db 9051 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9097

RESULT 35

PTR270072 10151 bp DNA PRI 05-NOV-1999

LOCUS Pan troglodytes troglodytes Xql3.3 non-coding region, C6.

DEFINITION AJ270072

ACCESSION AJ270072

VERSION AJ270072.1 GI:6273637

KEYWORDS non-coding region.

SOURCE chimpanzee.

ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 (bases 1 to 10151)
Kaessmann,H., Wiebe,V. and Paabo,S.
Extensive nuclear DNA sequence diversity among chimpanzees
Science 286 (5442), 1159-1162 (1999)

REFERENCE 2 (bases 1 to 10151)
Kaessmann,H.G.F.
Direct Submission
Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY

FEATURES
source
1. .10151
/organism="Pan troglodytes"
/sub_species="troglodytes"
/db_xref="taxon:9598"
/country="Gabon"
/chromosome="X"
/sex="male"
/map="Xql3.3"
/note="C4"
/note="non-coding DNA"
1. .10151
/note="non-coding DNA" 3223 a 2031 c 1818 g 3079 t

BASE COUNT 3223 a 2031 c 1818 g 3079 t

ORIGIN

Query Match 4.5%; Score 47; DB 9; Length 10151; Best Local Similarity 100.0%; Pred. No. 1.5e-14; Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacgggttcgccatgttgaccaggctggtctcgaaactc 696
|||||

Db 9051 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9097

RESULT 34

PTR270071 10151 bp DNA PRI 05-NOV-1999

LOCUS Pan troglodytes troglodytes Xql3.3 non-coding region, C8.

DEFINITION

ACCESSION AJ270071

VERSION AJ270071.1 GI:6273636

KEYWORDS non-coding region.

SOURCE chimpanzee.

ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 (bases 1 to 10151)
Kaessmann,H., Wiebe,V. and Paabo,S.
Extensive nuclear DNA sequence diversity among chimpanzees
Science 286 (5442), 1159-1162 (1999)

REFERENCE 2 (bases 1 to 10151)
Kaessmann,H.G.F.
Direct Submission
Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY

FEATURES
source
1. .10151
/organism="Pan troglodytes"
/sub_species="troglodytes"
/db_xref="taxon:9598"
/country="Gabon"
/chromosome="X"
/sex="male"
/map="Xql3.3"
/note="C8"
/note="non-coding DNA"
1. .10151
/note="non-coding DNA" 3221 a 2032 c 1818 g 3080 t

BASE COUNT 3221 a 2032 c 1818 g 3080 t

ORIGIN

Query Match 4.5%; Score 47; DB 9; Length 10151; Best Local Similarity 100.0%; Pred. No. 1.5e-14; Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacgggttcgccatgttgaccaggctggtctcgaaactc 696
|||||

Db 9051 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9097

RESULT 35

PTR270072 10151 bp DNA PRI 05-NOV-1999

LOCUS Pan troglodytes troglodytes Xql3.3 non-coding region, C6.

DEFINITION AJ270072

ACCESSION AJ270072

VERSION AJ270072.1 GI:6273637

KEYWORDS non-coding region.

SOURCE chimpanzee.

ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 (bases 1 to 10151)
Kaessmann,H., Wiebe,V. and Paabo,S.
Extensive nuclear DNA sequence diversity among chimpanzees
Science 286 (5442), 1159-1162 (1999)

REFERENCE 2 (bases 1 to 10151)
Kaessmann,H.G.F.
Direct Submission
Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY

FEATURES
source
1. .10151
/organism="Pan troglodytes"
/sub_species="troglodytes"
/db_xref="taxon:9598"
/country="Gabon"
/chromosome="X"
/sex="male"

```
misc_feature      /map="Xql3.3"
                  /note="C6"
                  /note="non-coding DNA"
BASE COUNT      3222 a 2032 c 1815 g 3082 t
ORIGIN

Query Match
Best Local Similarity 4.5%; Score 47; DB 9; Length 10151;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 650 tagtagagacggggttcgccatgttgaccaggctggtctcgaaactc 696
|||||
Db 9051 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9097

RESULT 36
PTR241092
LOCUS
DEFINITION Pan troglodytes subsp. verus Xql3.3 non-coding region, isolated
from West Africa.
ACCESSION AJ241092
VERSION AJ241092.1 GI:4775559
KEYWORDS non-coding region.
SOURCE Pan troglodytes verus.
ORGANISM Pan troglodytes verus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 10152)
AUTHORS Kaessmann,H., Heissig,F., von Haeseler,A. and Paabo,S.
TITLE DNA sequence variation in a non-coding region of low recombination
on the human X chromosome
JOURNAL Nat. Genet. 22 (1), 78-81 (1999)
MEDLINE 99251584
REFERENCE 2 (bases 1 to 10152)
AUTHORS Kaessmann,H.
TITLE Direct Submission
JOURNAL Submitted (22-MAR-1999) Kaessmann H, Max-Planck-Institute of
Evolutionary Anthropology Inselstrasse 22, 04103 Leipzig, FRG
FEATURES
source
Location/Qualifiers
1..10152
/organism="Pan troglodytes verus"
/sub_species="verus"
/isolate="from West Africa"
/db_xref="taxon:37012"
/chromosome="X"
/map="Xql3.3"
misc_feature
1..10152
/note="non-coding region"
BASE COUNT 3229 a 2028 c 1814 g 3081 t
ORIGIN

Query Match
Best Local Similarity 4.5%; Score 47; DB 9; Length 10152;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 650 tagtagagacggggttcgccatgttgaccaggctggtctcgaaactc 696
|||||
Db 9052 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9098

RESULT 37
PTR270061
LOCUS
DEFINITION Pan troglodytes troglodytes Xql3.3 non-coding region, (10154bp).
ACCESSION AJ270061
VERSION AJ270061.1 GI:6273626
KEYWORDS non-coding region.
SOURCE Pan troglodytes troglodytes.
ORGANISM Pan troglodytes troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
REFERENCE
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
TITLE 1 (bases 1 to 10154)
JOURNAL Kaessmann,H., Wiebe,V. and Paabo,S.
MEDLINE Extensive nuclear DNA sequence diversity among chimpanzees
20018332 Science 286 (5442), 1159-1162 (1999)
REFERENCE 2 (bases 1 to 10154)
AUTHORS Kaessmann,H.G.F.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
Max-planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY
FEATURES
source
Location/Qualifiers
1..10154
/organism="Pan troglodytes troglodytes"
/sub_species="troglodytes"
/db_xref="taxon:37011"
/chromosome="X"
/map="Xql3.3"
/sex="male"
/country="Gabon"
/note="C11"
misc_feature
1..10154
/note="non-coding DNA"
BASE COUNT 3228 a 2030 c 1815 g 3081 t
ORIGIN

Query Match
Best Local Similarity 4.5%; Score 47; DB 9; Length 10154;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 650 tagtagagacggggttcgccatgttgaccaggctggtctcgaaactc 696
|||||
Db 9054 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9100

RESULT 38
PTR270062
LOCUS
DEFINITION Pan troglodytes troglodytes Xql3.3 non-coding region, C12.
ACCESSION AJ270062
VERSION AJ270062.1 GI:6273627
KEYWORDS non-coding region.
SOURCE chimpanzee.
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 10154)
AUTHORS Kaessmann,H., Wiebe,V. and Paabo,S.
TITLE Extensive nuclear DNA sequence diversity among chimpanzees
JOURNAL Science 286 (5442), 1159-1162 (1999)
MEDLINE 20018332
REFERENCE 2 (bases 1 to 10154)
AUTHORS Kaessmann,H.G.F.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY
FEATURES
source
Location/Qualifiers
1..10154
/organism="Pan troglodytes"
/sub_species="troglodytes"
/db_xref="taxon:9598"
/country="Gabon"
/chromosome="X"
/sex="male"
/map="Xql3.3"
/note="C12"
misc_feature
1..10154
/note="non-coding DNA"
BASE COUNT 3227 a 2028 c 1814 g 3085 t
ORIGIN
```

Query Match 4.5%; Score 47; DB 9; Length 10154;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacggggttcgccatgttgaccaggctggtctcgaactc 696
|||||
Db 9054 TAGTAGAGACGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9100

RESULT 39
PTR270064 10154 bp DNA PRI 05-NOV-1999
LOCUS Pan troglodytes troglodytes Xql3.3 non-coding region, C10.
DEFINITION
ACCESSION AJ270064
VERSION AJ270064.1 GI:6273629
KEYWORDS non-coding region.
SOURCE chimpanzee.
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 10154)
AUTHORS Kaessmann,H., Wiebe,V. and Paabo,S.
TITLE Extensive nuclear DNA sequence diversity among chimpanzees
JOURNAL Science 286 (5442), 1159-1162 (1999)
MEDLINE 20018332
REFERENCE 2 (bases 1 to 10154)
AUTHORS Kaessmann,H.G.F.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY
Location/Qualifiers
1. .10154
/organism="Pan troglodytes"
/sub_species="troglodytes"
/db_xref="taxon:9598"
/country="Gabon"
/chromosome="X"
/sex="male"
/map="Xql3.3"
/note="C10"
misc_feature 1. .10154
/note="non-coding DNA"
BASE COUNT 3228 a 2030 c 1814 g 3082 t
ORIGIN

Query Match 4.5%; Score 47; DB 9; Length 10154;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacggggttcgccatgttgaccaggctggtctcgaactc 696
|||||
Db 9054 TAGTAGAGACGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9100

RESULT 40
PTR270067 10154 bp DNA PRI 05-NOV-1999
LOCUS Pan troglodytes troglodytes Xql3.3 non-coding region, C1.
DEFINITION
ACCESSION AJ270067
VERSION AJ270067.1 GI:6273632
KEYWORDS non-coding region.
SOURCE chimpanzee.
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 10154)
AUTHORS Kaessmann,H., Wiebe,V. and Paabo,S.
TITLE Extensive nuclear DNA sequence diversity among chimpanzees
JOURNAL Science 286 (5442), 1159-1162 (1999)

MEDLINE 20018332
REFERENCE 2 (bases 1 to 10154)
AUTHORS Kaessmann,H.G.F.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY
Location/Qualifiers
1. .10154
/organism="Pan troglodytes"
/sub_species="troglodytes"
/db_xref="taxon:9598"
/country="Gabon"
/chromosome="X"
/sex="male"
/map="Xql3.3"
/note="C1"
misc_feature 1. .10154
/note="non-coding DNA"
BASE COUNT 3227 a 2029 c 1815 g 3083 t
ORIGIN

Query Match 4.5%; Score 47; DB 9; Length 10154;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacggggttcgccatgttgaccaggctggtctcgaactc 696
|||||
Db 9054 TAGTAGAGACGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9100

RESULT 41
PTR270073 10154 bp DNA PRI 09-NOV-1999
LOCUS Pan troglodytes troglodytes Xql3.3 non-coding region, W13.
DEFINITION
ACCESSION AJ270073
VERSION AJ270073.1 GI:6318234
KEYWORDS non-coding region.
SOURCE chimpanzee.
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 10154)
AUTHORS Kaessmann,H., Wiebe,V. and Paabo,S.
TITLE Extensive nuclear DNA sequence diversity among chimpanzees
JOURNAL Science 286 (5442), 1159-1162 (1999)
MEDLINE 20018332
REFERENCE 2 (bases 1 to 10154)
AUTHORS Kaessmann,H.G.F.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY
Location/Qualifiers
1. .10154
/organism="Pan troglodytes"
/sub_species="verus"
/db_xref="taxon:9598"
/chromosome="X"
/sex="male"
/map="Xql3.3"
/note="W13"
misc_feature 1. .10154
/note="non-coding DNA"
BASE COUNT 3227 a 2028 c 1815 g 3084 t
ORIGIN

Query Match 4.5%; Score 47; DB 9; Length 10154;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacggggttcgccatgttgaccaggctggtctcgaactc 696
|||||
Db 9054 TACTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9100

RESULT 42
PTR270076

LOCUS PTR270076 10154 bp DNA PRI 08-NOV-1999
DEFINITION Pan troglodytes versus Xql3.3 non-coding region, W15.
ACCESSION AJ270076
VERSION AJ270076.1 GI:6318237
KEYWORDS non-coding region.
SOURCE chimpanzee.
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
AUTHORS Kaessmann,H., Wiebe,V. and Paabo,S.
TITLE Extensive nuclear DNA sequence diversity among chimpanzees
JOURNAL Science 286 (5442), 1159-1162 (1999)
MEDLINE 20018332
REFERENCE 2 (bases 1 to 10154)
AUTHORS Kaessmann,H.G.F.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY

FEATURES
source Location/Qualifiers
1..10154
/organism="Pan troglodytes"
/sub_species="verus"
/db_xref="taxon:9598"
/chromosome="X"
/sex="male"
/map="Xql3.3"
/note="W15"
misc_feature 3227 a 2027 c 1815 g 3085 t
/note="non-coding DNA"

BASE COUNT 3227 a 2027 c 1815 g 3085 t
ORIGIN

Query Match 4.5%; Score 47; DB 9; Length 10154;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacggggttcgccatgttgaccaggctggtctcgaactc 696
|||||
Db 9054 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9100

RESULT 43
PTR270077

LOCUS PTR270077 10154 bp DNA PRI 05-NOV-1999
DEFINITION Pan troglodytes versus Xql3.3 non-coding region, W5.
ACCESSION AJ270077
VERSION AJ270077.1 GI:6273700
KEYWORDS non-coding region.
SOURCE chimpanzee.
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
AUTHORS Kaessmann,H., Wiebe,V. and Paabo,S.
TITLE Extensive nuclear DNA sequence diversity among chimpanzees
JOURNAL Science 286 (5442), 1159-1162 (1999)
MEDLINE 20018332
REFERENCE 2 (bases 1 to 10154)
AUTHORS Kaessmann,H.G.F.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY

FEATURES
source Location/Qualifiers
1..10154
/organism="Pan troglodytes"
/sub_species="verus"
/db_xref="taxon:9598"
/country="Sierra Leone"
/chromosome="X"
/sex="male"
/map="Xql3.3"
/note="W5"
misc_feature 3229 a 2026 c 1814 g 3085 t
/note="non-coding DNA"

BASE COUNT 3229 a 2026 c 1814 g 3085 t
ORIGIN

Query Match 4.5%; Score 47; DB 9; Length 10154;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacggggttcgccatgttgaccaggctggtctcgaactc 696
|||||
Db 9054 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9100

RESULT 44
PTR270078

LOCUS PTR270078 10154 bp DNA PRI 05-NOV-1999
DEFINITION Pan troglodytes versus Xql3.3 non-coding region, W14.
ACCESSION AJ270078
VERSION AJ270078.1 GI:6273701
KEYWORDS non-coding region.
SOURCE chimpanzee.
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
AUTHORS Kaessmann,H., Wiebe,V. and Paabo,S.
TITLE Extensive nuclear DNA sequence diversity among chimpanzees
JOURNAL Science 286 (5442), 1159-1162 (1999)
MEDLINE 20018332
REFERENCE 2 (bases 1 to 10154)
AUTHORS Kaessmann,H.G.F.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY

FEATURES
source Location/Qualifiers
1..10154
/organism="Pan troglodytes"
/sub_species="verus"
/db_xref="taxon:9598"
/country="Sierra Leone"
/chromosome="X"
/sex="male"
/map="Xql3.3"
/note="W14"
misc_feature 3227 a 2028 c 1815 g 3084 t
/note="non-coding DNA"

BASE COUNT 3227 a 2028 c 1815 g 3084 t
ORIGIN

Query Match 4.5%; Score 47; DB 9; Length 10154;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacggggttcgccatgttgaccaggctggtctcgaactc 696
|||||
Db 9054 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9100

RESULT 45

PTR270079 10154 bp DNA PRI 05-NOV-1999
Pan troglodytes verus Xq13.3 non-coding region, W6.
AJ270079 GI:6273702
non-coding region.
chimpanzee.
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 (bases 1 to 10154)
Kaessmann,H., Wiebe,V. and Paabo,S.
Extensive nuclear DNA sequence diversity among chimpanzees
Science 286 (5442), 1159-1162 (1999)
20018332
2 (bases 1 to 10154)
Kaessmann,H.G.F.
Direct Submission
Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY
Location/Qualifiers
1..10154
/organism="Pan troglodytes"
/sub_species="verus"
/db_xref="taxon:9598"
/country="Sierra Leone"
/chromosome="X"
/sex="male"
/map="Xq13.3"
/note="W6"
misc_feature 1..10154
/note="non-coding DNA"
BASE COUNT 3228 a 2028 c 1814 g 3084 t
ORIGIN
Query Match 4.5%; Score 47; DB 9; Length 10154;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 650 tagtagagacgggttcgccatgttgaccaggctggtctcgaaactc 696
|||||
Db 9054 TAGTAGAGACGGGTTTCGCCATGTGACCAGGCTGGTCTCGAACTC 9100
RESULT 46
PTR270080 10154 bp DNA PRI 05-NOV-1999
Pan troglodytes verus Xq13.3 non-coding region, W12.
AJ270080 GI:6273703
non-coding region.
chimpanzee.
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 (bases 1 to 10154)
Kaessmann,H., Wiebe,V. and Paabo,S.
Extensive nuclear DNA sequence diversity among chimpanzees
Science 286 (5442), 1159-1162 (1999)
20018332
2 (bases 1 to 10154)
Kaessmann,H.G.F.
Direct Submission
Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY
Location/Qualifiers
1..10154
/organism="Pan troglodytes"
/sub_species="verus"
/db_xref="taxon:9598"
source

/country="Sierra Leone"
/chromosome="X"
/sex="male"
/map="Xq13.3"
/note="W12"
misc_feature 1..10154
/note="non-coding DNA"
BASE COUNT 3223 a 2030 c 1816 g 3085 t
ORIGIN
Query Match 4.5%; Score 47; DB 9; Length 10154;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 650 tagtagagacgggttcgccatgttgaccaggctggtctcgaaactc 696
|||||
Db 9054 TAGTAGAGACGGGTTTCGCCATGTGACCAGGCTGGTCTCGAACTC 9100
RESULT 47
PTR270081 10154 bp DNA PRI 05-NOV-1999
Pan troglodytes verus Xq13.3 non-coding region, W10.
AJ270081 GI:6273704
non-coding region.
chimpanzee.
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 (bases 1 to 10154)
Kaessmann,H., Wiebe,V. and Paabo,S.
Extensive nuclear DNA sequence diversity among chimpanzees
Science 286 (5442), 1159-1162 (1999)
20018332
2 (bases 1 to 10154)
Kaessmann,H.G.F.
Direct Submission
Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY
Location/Qualifiers
1..10154
/organism="Pan troglodytes"
/sub_species="verus"
/db_xref="taxon:9598"
/country="Sierra Leone"
/chromosome="X"
/sex="male"
/map="Xq13.3"
/note="W10"
misc_feature 1..10154
/note="non-coding DNA"
BASE COUNT 3229 a 2028 c 1814 g 3083 t
ORIGIN
Query Match 4.5%; Score 47; DB 9; Length 10154;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 650 tagtagagacgggttcgccatgttgaccaggctggtctcgaaactc 696
|||||
Db 9054 TAGTAGAGACGGGTTTCGCCATGTGACCAGGCTGGTCTCGAACTC 9100
RESULT 48
PTR270082 10154 bp DNA PRI 05-NOV-1999
Pan troglodytes verus Xq13.3 non-coding region, W1.
AJ270082 GI:6273705

KEYWORDS non-coding region.
SOURCE chimpanzee.
ORGANISM Pan troglodytes
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 10154)
AUTHORS Kaessmann,H., Wiebe,V. and Paabo,S.
TITLE Extensive nuclear DNA sequence diversity among chimpanzees
JOURNAL Science 286 (5442), 1159-1162 (1999)
MEDLINE 20018332
REFERENCE 2 (bases 1 to 10154)
AUTHORS Kaessmann,H.G.F.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY
FEATURES source
Location/Qualifiers
1. .10154
/organism="Pan troglodytes"
/sub_species="verus"
/db_xref="taxon:9598"
/country="Sierra Leone"
/chromosome="X"
/sex="male"
/map="Xq13.3"
/note="W1"
misc_feature 1. .10154
/note="non-coding DNA"
BASE COUNT 3227 a 2030 c 1815 g 3082 t
ORIGIN
Query Match 4.5%; Score 47; DB 9; Length 10154;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 650 tagtagagacggggttcgccatgttgaccaggctgggtctcgaactc 696
|||||
Db 9054 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9100
RESULT 49
PTR270083
LOCUS PTR270083 10154 bp DNA PRI 05-NOV-1999
DEFINITION Pan troglodytes verus Xql3.3 non-coding region, w8.
ACCESSION AJ270083
VERSION AJ270083.1 GI:6273706
KEYWORDS non-coding region.
SOURCE chimpanzee.
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 10154)
AUTHORS Kaessmann,H., Wiebe,V. and Paabo,S.
TITLE Extensive nuclear DNA sequence diversity among chimpanzees
JOURNAL Science 286 (5442), 1159-1162 (1999)
MEDLINE 20018332
REFERENCE 2 (bases 1 to 10154)
AUTHORS Kaessmann,H.G.F.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY
FEATURES source
Location/Qualifiers
1. .10154
/organism="Pan troglodytes"
/sub_species="verus"
/db_xref="taxon:9598"
/country="Sierra Leone"
/chromosome="X"
/sex="male"
/map="Xq13.3"
/note="W8"

misc_feature 1. .10154
/note="non-coding DNA"
BASE COUNT 3229 a 2028 c 1814 g 3083 t
ORIGIN
Query Match 4.5%; Score 47; DB 9; Length 10154;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 650 tagtagagacggggttcgccatgttgaccaggctgggtctcgaactc 696
|||||
Db 9054 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9100
RESULT 50
PTR270084
LOCUS PTR270084 10154 bp DNA PRI 05-NOV-1999
DEFINITION Pan troglodytes verus Xql3.3 non-coding region, w2.
ACCESSION AJ270084
VERSION AJ270084.1 GI:6273707
KEYWORDS non-coding region.
SOURCE chimpanzee.
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 10154)
AUTHORS Kaessmann,H., Wiebe,V. and Paabo,S.
TITLE Extensive nuclear DNA sequence diversity among chimpanzees
JOURNAL Science 286 (5442), 1159-1162 (1999)
MEDLINE 20018332
REFERENCE 2 (bases 1 to 10154)
AUTHORS Kaessmann,H.G.F.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY
FEATURES source
Location/Qualifiers
1. .10154
/organism="Pan troglodytes"
/sub_species="verus"
/db_xref="taxon:9598"
/country="Sierra Leone"
/chromosome="X"
/sex="male"
/map="Xq13.3"
/note="W2"
misc_feature 1. .10154
/note="non-coding DNA"
BASE COUNT 3228 a 2029 c 1815 g 3082 t
ORIGIN
Query Match 4.5%; Score 47; DB 9; Length 10154;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 650 tagtagagacggggttcgccatgttgaccaggctgggtctcgaactc 696
|||||
Db 9054 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9100
Search completed: May 2, 2002, 19:45:59
Job time: 9353 sec

us-09-671-050-9.oli20.rge

Fri May 3 11:02:09 2002

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2002, 11:01:25 ; Search time 40.16 Seconds
(without alignments)
640.026 Million cell updates/sec

Title: US-09-671-050-10
Perfect score: 1820
Sequence: 1 MEKYEKLAKTGESGYGVFK.....RKARNEGRNRRRQVLP LKS 347

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues 522463
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_1101.*
1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT.*
2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query %		DB ID	Description
	Score	Match		
1	1820	100.0	347 22 AAE00494	Human kinase #5.
2	1796	98.7	356 22 AAE00491	Human kinase #2.
3	1636	89.9	315 22 AAE00495	Human kinase #6.
4	1612	88.6	324 22 AAE00492	Human kinase #3.
5	1593	87.5	360 22 AAU03525	Human protein kina
6	1270.5	69.8	296 22 AAB65643	Novel protein kina
7	1123	61.7	247 22 AAB65642	Novel protein kina
8	941.5	51.7	566 21 AAY90724	Rabbit KKIAMRE kin
9	791	43.5	187 22 AAE00490	Human kinase #1.
10	791	43.5	198 22 AAE00493	Human kinase #4.
11	758.5	41.7	455 21 AAY70126	Human lost in leuk

Human protein kina
Rat lost in leukae
Rat lost in leukae
Novel protein kina
Human polypeptide
CDK2-green fluore
Human polypeptide
Homo sapiens Cdc2
Green fluorescent
Oryza sativa Cdc2
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Human protein kina
Zea mays protein f
Homo sapiens Cdc2
Zea mays protein f
Rattus norvegicus
Cyclin-dependent k
Pneumocystis carin
Protein involved i
Tomato cyclin-depe
Cyclin dependent p
Schizosaccharomyce
Fission yeast prot
Arabidopsis thalia
Emericella nidulan
CDK1 protein. Can
Candida albicans C
Candida cyclin dep
Ajellomyces capsul
Arabidopsis thalia
Arabidopsis thalia
Human cyclin depen

ALIGNMENTS

RESULT 1

AAE00494
ID AAE00494 standard; Protein; 347 AA.
XX
AC AAE00494;
XX
DT 19-JUN-2001 (first entry)
XX
DE Human kinase #5.
XX
KW Human; kinase; gene therapy; bioreactor; mental disorder;
KW biological disorder.
XX
OS Homo sapiens.
XX
PN WO200123579-A1.
XX
PD 05-APR-2001.
XX
PF 27-SEP-2000; 2000WO-US26621.
XX
PR 28-SEP-1999; 99US-0156511.
XX (LEXI-) LEXICON GENETICS INC.
PA Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
PI
XX
DR WPI; 2001-266166/27.
DR N-PSDB; AAD03816.
XX
PT New isolated human kinase polynucleotide useful for generating
PT antibodies, as reagents in diagnostic assays and for screening for
PT compounds useful for treating mental, biological or medical diseases
XX

PS Claim 2; Page 32-33; 38pp; English.

XX The present sequence is novel human protein
CC (NHP) known as human kinase. The human kinases share structural
CC similarity with animal kinases, more particularly serine or
CC threonine protein kinases. Human kinase cDNA is useful for the
CC detection of mutant human kinase for the diagnosis of disease,
CC and also as a therapeutic. It is useful for screening drugs
CC effective in the treatment of symptomatic or phenotypic
CC manifestations perturbing the normal function of NHP in the
CC body. The NHP nucleotide sequences are useful for generation of
CC antibodies, as reagents in diagnostic assays, for the
CC identification of other cellular gene products related to human
CC kinases, and as reagents in assays for screening compounds that
CC are useful for treating mental, biological or medical disorders.
CC NHP oligonucleotides are used as probes. The labelled NHP probes
CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.

XX Sequence 347 AA;

Query Match 100.0%; Score 1820; DB 22; Length 347;
Best Local Similarity 100.0%; Pred. No. 3.5e-187;
Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKYEKLAKTGE GSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1 mekyeklaktgegsygvvfkcrnktsgqvavkkfveseddppvkkialreirmkqlkh 60

QY 61 PNLVNLIEVFRRRKRKMHVFEYCDHTLLNELRNPNGVADGVKISVLWOTLQALNFCIH 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
61 pnlvnlievfrrrkrkmhlvfeycdhtllnelernpngvadgvkislwqlqalnfcih 120

QY 121 NCIHRDIKPENILITKOGIIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGD 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121 ncihrdikpenilitkogiikicdfgfaqilipgdaytdyvatrwyrapellvgdtgys 180

QY 181 SVDIWAIGCVFAELLTGQPLWPKSDVDQLYLIIRTLVETGFRHVDQAGLELLTSSDPPA 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
181 svidwaigcvfaellitgqplwpgksdvdllyliirtlvetgfrhvdqagllelltssdppa 240

QY 241 VASQSAGITGKLI PRHOSIFKSNFGFFHGISIPEPEMETLEEKFSDVHPVALNFMKGCLK 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
241 vasqsagitgkli prhqsifksgffhgisipepedmetleekfsdvhpvalnfmkgclk 300

QY 301 MNPDDRRLTCSQLLESSYFDSFQEAQIKRKARNEGRRRRRQV 347
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
301 mnpddrltcsqllessyfdsfgeaqikrkarnegrrrrrqv 347

RESULT 2
AAE00491
ID AAE00491 standard; Protein; 356 AA.
XX
AC AAE00491;
XX
DT 19-JUN-2001 (first entry)
XX
DE Human kinase #2.
XX
KW Human; kinase; gene therapy; bioreactor; mental disorder;
KW biological disorder.
XX
OS Homo sapiens.
XX

PN WO200123579-A1.
XX
PD 05-APR-2001.
XX
PF 27-SEP-2000; 2000WO-US26621.
XX
PR 28-SEP-1999; 99US-0156511.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
XX
DR WPI; 2001-266166/27.
DR N-PSDB; AAD03813.
XX
PT New isolated human kinase polynucleotide useful for generating
PT antibodies, as reagents in diagnostic assays and for screening for
PT compounds useful for treating mental, biological or medical diseases -
XX
PS Disclosure; Page 28-29; 38pp; English.
XX
CC The present sequence is novel human protein
CC (NHP) known as human kinase. The human kinases share structural
CC similarity with animal kinases, more particularly serine or
CC threonine protein kinases. Human kinase cDNA is useful for the
CC detection of mutant human kinase for the diagnosis of disease,
CC and also as a therapeutic. It is useful for screening drugs
CC effective in the treatment of symptomatic or phenotypic
CC manifestations perturbing the normal function of NHP in the
CC body. The NHP nucleotide sequences are useful for generation of
CC antibodies, as reagents in diagnostic assays, for the
CC identification of other cellular gene products related to human
CC kinases, and as reagents in assays for screening compounds that
CC are useful for treating mental, biological or medical disorders.
CC NHP oligonucleotides are used as probes. The labelled NHP probes
CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.

XX Sequence 356 AA;

Query Match 98.7%; Score 1796; DB 22; Length 356;
Best Local Similarity 100.0%; Pred. No. 1.4e-184;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKYEKLAKTGE GSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1 mekyeklaktgegsygvvfkcrnktsgqvavkkfveseddppvkkialreirmkqlkh 60

QY 61 PNLVNLIEVFRRRKRKMHVFEYCDHTLLNELRNPNGVADGVKISVLWOTLQALNFCIH 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
61 pnlvnlievfrrrkrkmhlvfeycdhtllnelernpngvadgvkislwqlqalnfcih 120

QY 121 NCIHRDIKPENILITKOGIIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGD 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121 ncihrdikpenilitkogiikicdfgfaqilipgdaytdyvatrwyrapellvgdtgys 180

QY 181 SVDIWAIGCVFAELLTGQPLWPKSDVDQLYLIIRTLVETGFRHVDQAGLELLTSSDPPA 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
181 svidwaigcvfaellitgqplwpgksdvdllyliirtlvetgfrhvdqagllelltssdppa 240

QY 241 VASQSAGITGKLI PRHOSIFKSNFGFFHGISIPEPEMETLEEKFSDVHPVALNFMKGCLK 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
241 vasqsagitgkli prhqsifksgffhgisipepedmetleekfsdvhpvalnfmkgclk 300

QY 301 MNPDDRRLTCSQLLESSYFDSFQEAQIKRKARNEGRRRRRQV 342

Db 301 mnpddrltcsqllessyfdsfgeaqikrkarnegrrrrgqv 342
|||||

RESULT 3
AAE00495
ID AAE00495 standard; Protein; 315 AA.

XX AAE00495;
XX 19-JUN-2001 (first entry)
XX Human kinase #6.

DE Human; kinase; gene therapy; bioreactor; mental disorder;
KW biological disorder.

XX Homo sapiens.
OS WO200123579-A1.

XX 05-APR-2001.
XX 27-SEP-2000; 2000WO-US26621.
XX 28-SEP-1999; 99US-0156511.
XX (LEXI-) LEXICON GENETICS INC.

XX Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
PI WPI; 2001-266166/27.
XX N-PSDB; AAD03817.

XX New isolated human kinase polynucleotide useful for generating
PT antibodies, as reagents in diagnostic assays and for screening for
PT compounds useful for treating mental, biological or medical diseases

PS Claim 3; Page 33-34; 38pp; English.

XX The present sequence is novel human protein
CC (NHP) known as human kinase. The human kinases share structural
CC similarity with animal kinases, more particularly serine or
CC threonine protein kinases. Human kinase cDNA is useful for the
CC detection of mutant human kinase for the diagnosis of disease,
CC and also as a therapeutic. It is useful for screening drugs
CC effective in the treatment of symptomatic or phenotypic
CC manifestations perturbing the normal function of NHP in the
CC body. The NHP nucleotide sequences are useful for generation of
CC antibodies, as reagents in diagnostic assays, for the
CC identification of other cellular gene products related to human
CC kinases, and as reagents in assays for screening compounds that
CC are useful for treating mental, biological or medical disorders.
CC NHP oligonucleotides are used as probes. The labelled NHP probes
CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.

XX Sequence 315 AA;

Query Match 89.9%; Score 1636; DB 22; Length 315;
Best Local Similarity 90.8%; Pred. No. 2e-167;
Matches 315; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

QY 1 MEKYEKLAKTGEGSYGVVFKCRNKTSGQVAVKRFVESEDDPVVKKIALREIRMLKQLKH 60
|||||

Db 1 mekyeklaktgegsygvvfkcrnktsgqvavvkfveseddpvkkialreirmlkqlkh 60
QY 61 PNLVNLIEVFRKRKMHLPFEYCDHTLLNELERNPNGVADGVKISVLWQTLQALNFCIH 120
Db 61 pnlvnlievfrkrkmlhvfeycdhtllnelernpnavgsviksvlwtlqalnfchih 120
QY 121 NCJHRDIKPENILITKOGIIKICDFGFAQILIPGDAYTDYVATRWYRAPPELLVGDYQGS 180
Db 121 ncjhrdikpenilitkqglikicdfgfaqilipgdaytdyvatrwyrappellvgdtqygs 180
QY 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRILVETGFRHVDQAGLELLTSSDPPA 240
Db 181 svdiwaigcvfaelltgqplwpgksdvdqlyliirtl----- 217
QY 241 VASQSAGITGKLIPRHQSIFKSNFGFFHGHSIPEPEMETLEEKFSVDVHPVALNFMKGCLK 300
Db 218 -----gkliprhqsifksngffhgisipepedmetleekfsdvhpvalnfmkgclk 268
QY 301 MNPDDRLTCSQLLESSYFDSFQEAQIKRKARNEGRNRRRQQVLPLKS 347
Db 269 mnpddrltcsqllessyfdsfgeaqikrkarnegrrrrqqvlpks 315

RESULT 4
AAE00492
ID AAE00492 standard; Protein; 324 AA.

XX AAE00492;
XX 19-JUN-2001 (first entry)
XX Human kinase #3.

XX Human; kinase; gene therapy; bioreactor; mental disorder;
KW biological disorder.

XX Homo sapiens.
XX WO200123579-A1.

XX 05-APR-2001.
XX 27-SEP-2000; 2000WO-US26621.

XX 28-SEP-1999; 99US-0156511.
XX (LEXI-) LEXICON GENETICS INC.

XX Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
PI WPI; 2001-266166/27.
XX N-PSDB; AAD03814.

XX New isolated human kinase polynucleotide useful for generating
PT antibodies, as reagents in diagnostic assays and for screening for
PT compounds useful for treating mental, biological or medical diseases

PS Disclosure; Page 30; 38pp; English.

XX The present sequence is novel human protein
CC (NHP) known as human kinase. The human kinases share structural
CC similarity with animal kinases, more particularly serine or
CC threonine protein kinases. Human kinase cDNA is useful for the
CC detection of mutant human kinase for the diagnosis of disease,
CC and also as a therapeutic. It is useful for screening drugs
CC effective in the treatment of symptomatic or phenotypic
CC manifestations perturbing the normal function of NHP in the
CC body. The NHP nucleotide sequences are useful for generation of
CC antibodies, as reagents in diagnostic assays, for the
CC identification of other cellular gene products related to human
CC kinases, and as reagents in assays for screening compounds that
CC are useful for treating mental, biological or medical disorders.
CC NHP oligonucleotides are used as probes. The labelled NHP probes

CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.
XX
XX
SQ Sequence 324 AA;

Query Match 88.6%; Score 1612; DB 22; Length 324;
Best Local Similarity 90.6%; Pred. No. 7.9e-165;
Matches 310; Conservative 0; Mismatches 0; Indels 32; Gaps 1;
QY 1 MEKYEKLAKTGESYGVVFKCRNKTSGQVVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db |||||
QY 1 mekyeklaktgegsygvvfkcrnktsgqvavkkfveseddpvkkialreirmlkqlkh 60
Db |||||
QY 61 PNLVNLIEVFRKRKMHLVFECYCDHTLLNELERNPVGADGVKSVLWOTLQALNFCIH 120
Db |||||
QY 61 pnlvnlievfrkrkmhlvfeycdhtllnelernpvgadviksvlwtlqalnfcih 120
Db |||||
QY 121 NCIHRIKPENILITKQGIKICDFGFAQILIPGDAYTDYVATRWYRAPPELLVGDYQGS 180
Db |||||
QY 121 ncihrdikpenilitkgikicdfgfaqilipdaytdyvatrwyrappellvgdtqygs 180
QY 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLVTETGRHVDQAGLELLTSSDPPA 240
Db |||||
QY 181 svidiwaigcvfaelltgplwpgksdvdqlyliirtl----- 217
QY 241 VASQSAGITGKLIPRHQSIFKSNGFHGISIPEPEMETLEEKFSVHPVALNFMKGCLK 300
Db |||||
QY 218 -----gkliprhqsifksngffhgisipepedmetleekfsdvhpvalnfmkgclk 268
QY 301 MNPDDLRLTCSQLLESYFDSFQEAQIKRKARNEGRRRRQQV 342
Db |||||
QY 269 mnpddrltcsqllessyfsdfsgaqlkrkarnegrnrrrqv 310

RESULT 5
AAU03525
ID AAU03525 standard; Protein; 360 AA.
XX
AC AAU03525;
DT 12-SEP-2001 (first entry)
XX
DE Human protein kinase #25.
XX
KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
KW metabolic disorder; immune related disease; neurological disorder;
KW neurodegenerative disorder; inflammatory disorder; infectious disease;
XX reproductive disorder.
OS Homo sapiens.
XX
PN WO200138503-A2.
XX
PD 31-MAY-2001.
XX
PF 22-NOV-2000; 2000WO-US32085.
XX
PR 24-NOV-1999; 99US-0167482.
XX
PA (SUGE-) SUGEN INC.
XX
PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
PI Flanagan P, Clary D;
XX
DR WPI; 2001-343950/36.

DR N-PSDB; AAS06725.
XX
PT Nucleic acids encoding human kinase polypeptides, useful for preventing
PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT neuronal-associated diseases, and microbial infections .
XX
PS Claim 7; Figure 2; 433pp; English.
XX
CC AAU03501-AAU03557 represent novel human protein kinases #1-57. The
CC novel protein kinases have been identified as members of the tyrosine
CC or serine/threonine kinase (PTK and STK) families. The polynucleotides
CC encoding protein kinases and the polypeptides may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate kinase expression. For example, they may be used to treat
CC cancers (especially cancers of haematopoietic origin), cardiovascular
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC immune related diseases (e.g. rheumatoid arthritis), neurological
CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC Additionally, polynucleotides encoding protein kinases may be
CC used for gene therapy and as DNA probes in diagnostic assays.
CC The protein kinase polypeptides may be used as antigens in the production
CC of antibodies against the protein kinases and in assays to identify
CC modulators of protein kinase expression and activity.
XX
SQ Sequence 360 AA;

Query Match 87.5%; Score 1593; DB 22; Length 360;
Best Local Similarity 85.2%; Pred. No. 1e-162;
Matches 311; Conservative 2; Mismatches 0; Indels 52; Gaps 3;
QY 1 MEKYEKLAKTGESYGVVFKCRNKTSGQVVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db |||||
QY 1 mekyeklaktgegsygvvfkcrnktsgqvavkkfveseddpvkkialreirmlkqlkh 60
Db |||||
QY 61 PNLVNLIEVFRKRKMHLVFECYCDHTLLNELERNPVGADGVKSVLWOTLQALNFCIH 120
Db |||||
QY 61 pnlvnlievfrkrkmhlvfeycdhtllnelernpvgadviksvlwtlqalnfcih 120
QY 121 NCIHRIKPENILITKQGIKICDFGFAQIL-----IPGDAYTDYVA 162
Db |||||
QY 121 ncihrdikpenilitkgikicdfgfaqilnsnhvgrldplldafavpgdaytdyva 180
QY 163 TRWYRAPPELLVGDYQGSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLVTETGF 222
Db |||||
QY 181 trwyrapellvgdtqygssvdiwaigcvfaelltgplwpgksdvdqlyliirtl----- 235
QY 223 RHVDQAGLELLTSSDPPAVASQSAGITGKLIPRHQSIFKSNGFHGISIPEPEMETLEE 282
Db |||||
QY 283 KFSDVHPVALNFMKGCLKMNPDDLRLTCSQLLESYFDSFQEAQIKRKARNEGRRRRQ-- 340
Db |||||
QY 269 kfsdvhpvalnfmkgclkmpddrltcsqllessyfsdfsgaqlkrkarnegrnrrrqgn 328
QY 341 QVLPL 345
Db |::|
Db 329 qlpl 333

RESULT 6
AAB65643
ID AAB65643 standard; Protein; 296 AA.
XX
AC AAB65643;
XX
DT 27-MAR-2001 (first entry)
XX
DE Novel protein kinase, SEQ ID NO: 170.
XX
KW Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;

Db 1 ngvadgviksvlwtlqalnfcnhncihrdikpenilitkgiikicdfgfaqilipgd 60

QY 156 AYTDYVATRWYRAPELLVGDTOYGSSVDIWAIGCVFAELLTGQPLWPKSDVDQLYLIIR 215
|||||

Db 61 aytdyvatrwyrapelldvgtqygssvdiwaigcvfaelltgqplwpgksdvdqlyliir 120
|||||

QY 216 TLVETGFRHVDAQGLELLTSSDPPAVASQSAGITGKLIPRHOSIFKSNGFFHGISIPEPE 275
|||

Db 121 tl-----gkliprhqsifkngffhgisipepe 148

QY 276 DMETLEEKFSVHPVALNFMKGCLKMNPDRLTCSQLLESSYFDSQEAQIKRKARNEGR 335
|||||

Db 149 dmetleekfsdvhpvalnfmkgclkmnpddrltcsqlessyfdsfqeaqikrkarnegr 208

QY 336 NRRRQ--QVLPL 345
|||||

Db 209 nrrrqngllpl 220

RESULT 8

AAAY90724

ID AAAY90724 standard; Protein; 566 AA.

XX AC AAAY90724;

XX DT 15-AUG-2000 (first entry)

XX DE Rabbit KKIAMRE kinase SEQ ID NO:4.

XX KW Rabbit; KKIAMRE kinase; learning-induced kinase; learning; memory;

XX KW cdc2-related kinase; brain; gene therapy; genetic disorder; detection;

XX KW identification.

XX OS Oryctolagus cuniculus.

XX PN WO200020567-A2.

XX PD 13-APR-2000.

XX PF 01-OCT-1999; 99WO-US23010.

XX PR 02-OCT-1998; 98US-0102906.

XX PA (UYSC-) UNIV SOUTHERN CALIFORNIA.

XX PI Thompson RF, Gomi H, Sun W;

XX DR WPI; 2000-328932/28.

XX DR N-PSDB; AAA29745.

XX PT Novel learning induced kinase polynucleotides and polypeptides, useful

XX PT for the analysis of learning and memory, and for gene therapy -

XX PS Claim 1; Fig 4; 64pp; English.

XX CC The present sequence represents a learning-induced kinase, designated

CC KKIAMRE kinase, which is isolated from rabbit brain tissue. KKIAMRE

CC kinase is a cdc2-related kinase. The KKIAMRE kinase polynucleotides can

CC be used to express recombinant protein for analysis, characterisation or

CC therapeutic use, as markers for tissues in which the protein is

CC preferentially expressed, as molecular weight markers on Southern gels,

CC as chromosome markers or tags, to compare endogenous DNA sequences in

CC patients to identify potential genetic disorders, as probes to hybridise

CC and discover novel related sequences, as a source of PCR primers, and as

CC an antigen to induce anti-DNA antibodies. The polypeptides can be used

CC in assay to discover biological activity, to raise antibodies, as tissue

CC markers, and to isolate correlative receptors or ligands. The

CC polynucleotides may also be used for gene therapy for the treatment of

XX disorders which are mediated by KKIAMRE kinase.

SQ Sequence 566 AA;

Query Match 51.7%; Score 941.5; DB 21; Length 566;
Best Local Similarity 49.7%; Pred. No. 3e-92;
Matches 177; Conservative 67; Mismatches 65; Indels 47; Gaps 6;

QY 1 MEKYEKLAKTGEGSYGVVFKCRNKTSGQVAVKVKFESEDDPVWKKIALREIRMLKQLKH 60
|||||

Db 1 mekyenlglvgegsygmvmkcrnkdsgrivaikkflesddkmvkkiamreikilkqlrh 60
|||||

QY 61 PNLVNLIEVFRKRKRKMHLVFEYCDHTLLNELERNPNGVADGVIKSVLWQTLQALNFCIH 120
|||||

Db 61 enlvnlievckkkkrwylvfefvdhtildlelfpngiddqvqkylfqiingigfchsh 120
|||||

QY 121 NCIHRIKIPENILITKQIIKICDFGFAQIL-IPGDAVTDYVATRWYRAPELLVGDTOYG 179
|||||

Db 121 niihrdikpenilvsqsgvklcdfgfartlaapgevytdyvatrwyrapelldvdkyg 180
|||||

QY 180 SSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLVETGFRHVDAQGLELLTSSDPP 239
:|||||

Db 181 kavdvwaigclvtemlmgeplfpgdsdidqlylmrcf----- 218

QY 240 AVASQSAGITGKLIPRHQSIFKSNGFFHGISIPEPEDMETLEEKFSVHPVALNFMKGCL 299
| |||||

Db 219 -----gnliprhqelfyknppvfvagvrlpeikeseplerypklsevvldlakkcl 268
| |||||

QY 300 KMNPDRLTCSQLLESSYF--DSF-----OEAQIK--RKARN-----EGRNRRRQQ 341
:|||||

Db 269 hvdprkrfcaellhhdffqmdgfaerfsqelgmkvqkdarnislslksksqrkkek 324
:|||||

RESULT 9

AAE00490

ID AAE00490 standard; Protein; 187 AA.

XX AC AAE00490;

XX DT 19-JUN-2001 (first entry)

XX DE Human kinase #1.

XX KW Human; kinase; gene therapy; bioreactor; mental disorder;

XX KW biological disorder.

XX OS Homo sapiens.

XX PN WO200123579-A1.

XX PD 05-APR-2001.

XX PF 27-SEP-2000; 2000WO-US26621.

XX PR 28-SEP-1999; 99US-0156511.

XX PA (LEXI-) LEXICON GENETICS INC.

XX PI Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;

XX DR WPI; 2001-266166/27.

XX DR N-PSDB; AAD03812, AAD03818.

PT New isolated human kinase polynucleotide useful for generating

PT antibodies, as reagents in diagnostic assays and for screening for

PT compounds useful for treating mental, biological or medical diseases -

PS Disclosure; Page 27-28; 38pp; English.

XX CC The present sequence is novel human protein (NHP) known

CC as human kinase. The human kinases share structural

CC similarity with animal kinases, more particularly serine or

CC threonine protein kinases. Human kinase cDNA is useful for the

CC detection of mutant human kinase for the diagnosis of disease,

CC and also as a therapeutic. It is useful for screening drugs

CC effective in the treatment of symptomatic or phenotypic

CC manifestations perturbing the normal function of NHP in the

CC body. The NHP nucleotide sequences are useful for generation of
CC antibodies, as reagents in diagnostic assays, for the
CC identification of other cellular gene products related to human
CC kinases, and as reagents in assays for screening compounds that
CC are useful for treating mental, biological or medical disorders.
CC NHP oligonucleotides are used as probes. The labelled NHP probes
CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.
XX
SQ Sequence 187 AA;

Query Match 43.5%; Score 791; DB 22; Length 187;
Best Local Similarity 100.0%; Pred. No. 8.5e-77;
Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 mekyeklaktgeysygvvfkcrnktsgqvavkkfveseddpvkkialreirmlkqlkh 60
QY 61 PNLVNLIEVFRRRKRKMHVFEYCDHTLLNELEPNPGVADGVIKSVLWQTLQALNFCIH 120
Db 61 pnlvnlievfrkrkrmhlvfeycdhtllnelernpngvadgviksvlwqtlgalnfchih 120
QY 121 NCIHREDIKPENILITKQGIKICDFGFAQIL 151
Db 121 ncihrdikpenilitkgiikicdfgfaqil 151

RESULT 10
AAE00493
ID AAE00493 standard; Protein; 198 AA.
XX
AC AAE00493;
XX 19-JUN-2001 (first entry)
XX Human kinase #4.
DE Human; kinase; gene therapy; bioreactor; mental disorder;
KW biological disorder.
XX Homo sapiens.
XX WO200123579-A1.
XX 05-APR-2001.
XX 27-SEP-2000; 2000WO-US26621.
XX 28-SEP-1999; 99US-0156511.
XX (LEXI-) LEXICON GENETICS INC.
XX Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
PI WPI; 2001-266166/27.
XX N-PSDB; AAD03815.
DR New isolated human kinase polynucleotide useful for generating
XX antibodies, as reagents in diagnostic assays and for screening for
PT compounds useful for treating mental, biological or medical diseases
PT
XX Disclosure; Page 31; 38pp; English.
PS The present sequence is novel human protein
XX
CC

CC (NHP) known as human kinase. The human kinases share structural
CC similarity with animal kinases, more particularly serine or
CC threonine protein kinases. Human kinase cDNA is useful for the
CC detection of mutant human kinase for the diagnosis of disease,
CC and also as a therapeutic. It is useful for screening drugs
CC effective in the treatment of symptomatic or phenotypic
CC manifestations perturbing the normal function of NHP in the
CC body. The NHP nucleotide sequences are useful for generation of
CC antibodies, as reagents in diagnostic assays, for the
CC identification of other cellular gene products related to human
CC kinases, and as reagents in assays for screening compounds that
CC are useful for treating mental, biological or medical disorders.
CC NHP oligonucleotides are used as probes. The labelled NHP probes
CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.
XX
SQ Sequence 198 AA;

Query Match 43.5%; Score 791; DB 22; Length 198;
Best Local Similarity 100.0%; Pred. No. 9.3e-77;
Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 mekyeklaktgeysygvvfkcrnktsgqvavkkfveseddpvkkialreirmlkqlkh 60
QY 61 PNLVNLIEVFRRRKRKMHVFEYCDHTLLNELEPNPGVADGVIKSVLWQTLQALNFCIH 120
Db 61 pnlvnlievfrkrkrmhlvfeycdhtllnelernpngvadgviksvlwqtlqalnfcchih 120
QY 121 NCIHREDIKPENILITKQGIKICDFGFAQIL 151
Db 121 ncihrdikpenilitkgiikicdfgfaqil 151

RESULT 11
AAAY70126
ID AAAY70126 standard; Protein; 455 AA.
XX
AC AAAY70126;
XX 06-JUN-2000 (first entry)
DT Human lost in leukaemia kinase (LLK).
XX
DE Lost in leukaemia kinase; LLK; cdc2-related kinase; human; leukaemia;
XX MAPK; mitogen activated protein kinase; tumour suppressor; diagnosis;
KW chromosome 5q31.1; cytostatic; mitotic index; treatment; prevention;
KW gene therapy.
XX Homo sapiens.
OS
XX Location/Qualifiers
FH Key 11..16
FT Binding-site /label= ATP_binding_site
FT Binding-site 44..50
FT /label= Cyclin_binding_motif
FT Region 125..130
FT /label= Consensus_sequence
FT /note= "Serine/threonine specific kinase sequence"
FT Domain 158..160
FT /label= MAP_kinase_activation_motif
FT Region 162..167
FT /label= Consensus_sequence
FT /note= "Serine/threonine specific kinase sequence"

Db 218 -----gnlsphlqnifskspifagvlpqvqhpknarkkypknglladivhac1 267
QY 300 KMNPDRLTCSQLLESSYF--DSFQE---AQIKRKARNEGR 335
Db 268 qidpadrisssdlhheyftrdgdgfielkfmpe1kak1l1q1eak 308

RESULT 13
AAY70125
ID AAY70125 standard; Protein; 457 AA.
XX AAY70125;
XX 06-JUN-2000 (first entry)
DT Rat lost in leukaemia kinase beta isoform (LLK beta).
DE
XX
XX Lost in leukaemia kinase; LLK beta; rat; cdc2-related kinase; leukaemia;
KW MAPK; mitogen activated protein kinase; tumour suppressor; diagnosis;
KW cytostatic; mitotic index; treatment; prevention; gene therapy.
XX
OS Rattus sp.
XX
XX
FH Key Location/Qualifiers
FT Region 14..15
FT /label= MAP_kinase-like_regulatory_region
FT /note= "SY duplex"
FT Binding-site 44..50
FT /label= Cyclin_binding_motif
FT Region 125..130
FT /label= Consensus_sequence
FT /note= "Serine/threonine specific kinase sequence"
FT Region 158..160
FT /label= MAP_kinase-like_regulatory_region
FT Region 162..167
FT /label= Consensus_sequence
FT /note= "Serine/threonine specific kinase sequence"
XX WO200012719-A1.
XX
XX 09-MAR-2000.
XX
XX 31-AUG-1999; 99WO-CA00794.
XX
XX 31-AUG-1998; 98CA-2243784.
PR 20-NOV-1998; 98CA-2251249.
XX
XX (ONTA-) ONTARIO CANCER INST.
PA
XX Zanke B, Haq R, Randall S, Midmer M;
PI
XX
XX WPI; 2000-237880/20.
DR N-PSDB; AAZ51207.
XX
XX Isolated polynucleotide encoding a lost in leukemia kinase (LLK)
PT protein, useful for treatment, diagnosis and prevention of leukemia -
PT
XX
XX Claim 15; Page 48; 69pp; English.
PS
XX
XX The present amino acid sequence is the beta isoform of rat lost in
CC leukaemia kinase (LLK beta), isolated from rat jejunum cDNA library. It
CC is expressed strongly in muscle, heart, liver, brain and in tissues which
CC have a very low mitotic index. Rat LLK beta protein is closely related to
CC cdc2-related kinases, that are putative tumour suppressors and to mitogen
CC activated protein kinases (MAPKs). This sequence has cytostatic activity.
CC It is useful for the treatment, diagnosis and prevention of acute
CC leukaemia and is also used in gene therapy.
XX
SQ Sequence 457 AA;

Query Match

40.5%; Score 737.5; DB 21; Length 457;

Best Local Similarity 42.5%; Pred. No. 2e-70; Gaps 5;
Matches 145; Conservative 65; Mismatches 92; Indels 39;
QY 1 MEKYEKLAKTGEESYGWVFKCRNKTSQGVAVAKKFESEDDPVVKIALREIRMLKQLKH 60
Db 1 memyetlgkvgegsygtvmckkhdtdgrivaikfiye-kpeksvnkiatreiblkqfrh 59
QY 61 PNLVNLIEVFRKRKRMHLVFEYCDHTLLNELEPNPVGADGVIKSVLWQTLQALNFCIH 120
Db 60 enlvnlievfrqkkih1lvfefidhtvldelqhychglesrlrky1fqilraiey1hnn 119
QY 121 NCIHRDIKPENILITKGIKICDFGFAQIL-IPGDAYTDVATRWYRAPELLVGDQYG 179
Db 120 nihrdikpenilvsgitkklcdfgfartlaapgdvytdyvatrwyrapelv1kdttyg 179
QY 180 SSVDIWAIGCVFAELLTGQPLWPGKSDVDQYLIIRTLVETGFRHVDQAGLELLTSSDPP 239
Db 180 kpvdiwalgcmiematgnpylpssaldllhkv1kv----- 217
QY 240 AVASQSAGITGKLIPRHQSIFKSNGFHGISIPEPEDMETLEEKFSVDVHPVALNFMKGCL 299
Db 218 -----gnltphlhnifskspifagvlpqvqhpknarkkypknglladivhac1 267
QY 300 KMNPDRLTCSQLLESSYF--DSFQE---AQIKRKARNEGR 335
Db 268 qidpaerisstdllhhd1yftdrgdgiekfielrak1l1q1eak 308

RESULT 14

AAY70124
ID AAY70124 standard; Protein; 505 AA.
XX
AC AAY70124;
XX 06-JUN-2000 (first entry)
DT Rat lost in leukaemia kinase alpha isoform (LLK alpha).
DE
XX Lost in leukaemia kinase; LLK alpha; rat; cdc2-related kinase; leukaemia;
KW MAPK; mitogen activated protein kinase; tumour suppressor; diagnosis;
KW cytostatic; mitotic index; treatment; prevention; gene therapy.
XX
OS Rattus sp.
XX
XX
FH Key Location/Qualifiers
FT Region 14..15
FT /label= MAP_kinase-like_regulatory_region
FT /note= "SY duplex"
FT Binding-site 44..50
FT /label= Cyclin_binding_motif
FT Region 125..130
FT /label= Consensus_sequence
FT /note= "Serine/threonine specific kinase sequence"
FT Region 158..160
FT /label= MAP_kinase-like_regulatory_region
FT Region 162..167
FT /label= Consensus_sequence
FT /note= "Serine/threonine specific kinase sequence"
FT Domain 473..477
FT /label= Nuclear_localisation_motif
FT Region 480..493
FT /label= PEST_sequence
XX WO200012719-A1.
PN 09-MAR-2000.
XX 31-AUG-1999; 99WO-CA00794.
XX 31-AUG-1998; 98CA-2243784.
PR 20-NOV-1998; 98CA-2251249.
XX
XX (ONTA-) ONTARIO CANCER INST.

us-09-671-050-10.rag

Fri May 3 11:22:24 2002

Search completed: May 3, 2002, 11:02:15
Job time: 50 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2002, 11:01:26 ; Search time 26.4 Seconds
(without alignments)
1001.234 Million cell updates/sec

Title: US-09-671-050-10
Perfect score: 1820
Sequence: 1 MEKYEKLAKTGEYSYGVWFK.....RKARNEGRNRRRQQVLP LKS 347

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1157.5	63.6	358	2 S23383	protein kinase (EC
2	1149.5	63.2	376	2 S22745	serine/threonine p
3	565	31.0	305	2 S23382	protein kinase (EC
4	563	30.9	302	1 I50474	protein kinase (EC
5	558.5	30.7	292	2 S40021	protein kinase (EC
6	555	30.5	297	2 A37871	protein kinase (EC
7	554	30.4	298	2 A41227	protein kinase (EC
8	552.5	30.4	294	2 B40444	protein kinase (EC
9	551.5	30.3	294	2 S23095	protein kinase (EC
10	551.5	30.3	294	2 S22440	protein kinase (EC
11	546.5	30.0	294	1 A40444	protein kinase (EC
12	546.5	30.0	294	2 T49271	CELL DIVISION CONT
13	545	29.9	297	2 A29539	protein kinase (EC
14	545	29.9	298	1 A44878	protein kinase (EC
15	542	29.8	297	2 I45977	cyclin-dependent k
16	541	29.7	297	1 S24913	protein kinase (EC
17	540.5	29.7	297	1 S12009	protein kinase cdc
18	540.5	29.7	302	2 B44349	protein kinase (EC
19	540	29.7	302	1 A44349	protein kinase (EC
20	539	29.6	303	1 S06011	protein kinase (EC
21	537	29.5	288	1 S42566	protein kinase (EC
22	536.5	29.5	311	2 S36619	protein kinase (EC
23	535.5	29.4	294	1 J02243	protein kinase (EC
24	535.5	29.4	294	1 S57928	protein kinase (EC
25	535.5	29.4	294	1 S42049	protein kinase (EC
26	532	29.2	301	1 S19209	protein kinase (EC
27	531	29.2	292	1 A46365	protein kinase (EC
28	530.5	29.1	294	1 S31332	protein kinase (EC
29	529	29.1	292	1 I49592	protein kinase (EC

30	529	29.1	292	1 A45091	protein kinase (EC
31	527.5	29.0	308	1 S53538	protein kinase (EC
32	527	29.0	297	1 A36074	protein kinase (EC
33	527	29.0	301	1 S42101	protein kinase (EC
34	526.5	28.9	291	2 S23386	protein kinase (EC
35	526	28.9	293	2 JE0374	cyclin-dependent k
36	526	28.9	294	2 S51008	protein kinase (EC
37	526	28.9	296	2 S24386	protein kinase (EC
38	524	28.8	302	1 OKBY85	protein kinase PHO
39	521.5	28.7	302	2 T17115	protein kinase cdc
40	519.5	28.5	346	2 A54820	CDK-activating pro
41	518	28.5	292	2 S22441	protein kinase (EC
42	517.5	28.4	346	2 A56231	MO15/CDK-activatin
43	516.5	28.4	314	1 S12007	protein kinase (EC
44	514	28.2	293	2 T02922	protein kinase (EC
45	514	28.2	346	1 I78840	protein kinase (EC

ALIGNMENTS

RESULT 1
S23383
protein kinase (EC 2.7.1.37) cdc2-related KKIALRE - human
C;Species: Homo sapiens (man)
C;Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 18-Jun-1999
C;Accession: S23383; S22744
R;Meyerson, M.; Enders, G.H.; Wu, C.L.; Su, L.K.; Gorka, C.; Nelson, C.; Harlow, E.;
EMBO J. 11, 2909-2917, 1992
A;Title: A family of human cdc2-related protein kinases.
A;Reference number: S23382; MUID:92347325
A;Accession: S23383
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-358 <MEY>
A;Cross-references: EMBL:X66358; NID:g36614; PIDN:CAA47002.1; PID:g36615
C;Superfamily: Kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
F;3-278/Domain: protein kinase homology <KIN>
F;11-19/Region: protein kinase ATP-binding motif
F;34,52,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 63.6%; Score 1157.5; DB 2; Length 358;
Best Local Similarity 64.7%; Pred. No. 2.3e-46;
Matches 211; Conservative 41; Mismatches 41; Indels 33; Gaps 2;

QY	1	MEKYEKLAKTGEYSYGVVFKCRNKTSGQVAVVKKFVESEDDPVVKKIALREIRMLKQLKH	60
Db	2	MEKYEKIGKIGEGSYGVVFKCRNRTGQIVATKFKFLESEDDPVVKKIALREIRMLKQLKH	61
QY	61	PNLVNLIIEVFRKRKMHVFEYCDHTLLNELEPNNGVADGVKSVLWQTLQALNFCIH	120
Db	62	PNLVNLLIEVFRKRRLHLVFEYCDHTVLHELDYQYGVPEHLVKSITWQTLQAVNFCHKH	121
QY	121	NCIHRDIKPENILITKQGIKICDFGFAQILI-PGDAYTDYVATRWYRAPPELLVGDYQG	179
Db	122	NCIHRDVKPENILITKHSVIKLCDFGARLLTGPSDYITDYVATRWYRSPPELLVGDYQG	181
QY	180	SSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLVETGFRHVDQAGLELLTSSDPP	239
Db	182	PPVDWAIGCVFAELLSGVPLWPGKSDVDQLYLIRKTL-----	219
QY	240	AVASQSAGITGKLIPRHQSIFKSNGFHGISIPEDMETLEEFSDVHPVALNFMKGCL	299
Db	220	-----GDLIPRHQQVFSTNQYFSGVKIPDPEDMEPLELKFNPISYPALGLLKGL	269
QY	300	KMNPDRRLTCSQLLESSYFDSFQEAQ	325
Db	270	HMDPTERLTCEQLLHHPYFENIREIE	295

RESULT 2

S22745
serine/threonine protein kinase KKIALRE (EC 2.7.1.1) - human
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 19-Dec-1997
C:Accession: S22745
R:Meyerson, M.L.
submitted to the EMBL Data Library, May 1992
A:Reference number: S22743
A:Accession: S22745
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-376 <MEY>
A:Cross-references: EMBL:X66359
C:Genetics:
A:Introns: 152/3; 170/3
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
F:3-296/Domain: protein kinase homology <KIN>
F:11-19/Region: protein kinase ATP-binding motif

Query Match 63.2%; Score 1149.5; DB 2; Length 376;
Best Local Similarity 61.3%; Pred. No. 5.4e-46;
Matches 211; Conservative 41; Mismatches 41; Indels 51; Gaps 2;

QY 1 MEKYEKLAKTGEISYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 2 MEKYEKIGKIGESYGVVFKCRNRTGQIVAIKKFLESEDDPVVKKIALREIRMLKQLKH 61

QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEPNMGVADGVTKSVLWQTLQALNFC 120
Db 62 PNLVNLLEVFRKRRLHLVFEYCDHTVLHELDRIYQGVPEHLVKSIITWQTLQAVNFKKH 121

QY 121 NCIHRIKIPENILITKQGIKICDFGFAQILI-----PGDAYTDYV 161
Db 122 NCIHDRVKPENILITKHSVIKLCDFGARLLEFPQSAAVCPKCSITGTXGPSDYVDYV 181

QY 162 ATRWYRAPPELLVGDYQGSVDIWAIGCVFAELLTGQPLWPKSDVDQLYLIIRTLVETG 221
Db 182 ATRWYRSPPELLVGDYQGPVVDVWAIGCVFAELLSGVPLWPKSDVDQLYLIRKTL--- 237

QY 222 FRHVDQAGLELLTSSDPPAVASQSAGITGKLIPRHQSIFKSNFGFFHGISIPEPEDMETLE 281
Db 238 -----GDLIPRHQQVFSTNQYFSGVKIPDPEDMEPLE 269

QY 282 EKFSDVHPVALNFMKGCLKMNPDRDLTCSQLLESSYFDSFQEAQ 325
Db 270 LKFPNISYPALGLLKGLHMDPTERTLCEQLLHHPYFENIREIE 313

RESULT 3
S23382
protein kinase (EC 2.7.1.37) cdk3 - human
C:Species: Homo sapiens (man)
C:Date: 18-Jun-1993 #sequence_revision 18-Jun-1993 #text_change 18-Jun-1999
C:Accession: S23382; S22743
R:Meyerson, M.; Enders, G.H.; Wu, C.L.; Su, L.K.; Gorka, C.; Nelson, C.; Harlow, E.; Tsai
EMBO J. 11, 2909-2917, 1992
A:Title: A family of human cdc2-related transforming protein kinases.
A:Reference number: S23382; MUID:92347325
A:Accession: S23382
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-305 <MEY>
A:Cross-references: EMBL:X66357; NID:g36612; PIDN:CAA47001.1; PID:g36613
C:Genetics:
A:Gene: GDB:CDK3
A:Cross-references: GDB:283456
A:Map position: 12q13-12q13
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein kinase
F:2-255/Domain: protein kinase homology <KIN>
F:10-18/Region: protein kinase ATP-binding motif

F:33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 31.0%; Score 565; DB 2; Length 305;
Best Local Similarity 36.8%; Pred. No. 1.6e-19;
Matches 124; Conservative 66; Mismatches 99; Indels 48; Gaps 9;

QY 1 MEKYEKLAKTGEISYGVVFKCRNKTSGQVAVKKF---VESEDDPVVKKIALREIRMLKQ 57
Db 1 MDMFQKVEKIGEGTYGVVYKAKNRETGQVAKIRLDLEMEGVP---STAIREISLLKE 57

QY 58 LKHPNLVNLIEVFRKRKMHVFEYCDHTLLNELEPNMG-VADGVTKSVLWQTLQALNF 116
Db 58 LKHPNIVRLLDVVHNERKLYLVPEFLSQDLKKYMDSTPGSELPLHLIKSYLQQLQGVSF 117

QY 117 CHIHNCIHRDIKIPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPPELLVGD 175
Db 118 CHSHRVIHRDLKPNQLLNLINELGAIKLADFGLARAFGVPLRTYTHEVVTLWYRAPEILLGS 177

QY 176 TQYGSVDIWAIGCVFAELLTGQPLWPKSDVDQLYLIIRTLVETGFRHVDQAGLELLTS 235
Db 178 KFYTTAVDIWSIGCIFAEMVTRKALFPGDSEIDQLFRIFRML----- 219

QY 236 SDPPAVASQSA--GITGKLIPRHQSIFKSNFGFFHGISIPEPE-DMETLEEKFSDVHPVAL 292
Db 220 ----GTPSEDITWPGVT--QLPDYKGSF-----PKWTRKGLEEIVPNLEPEGR 260

QY 293 NFMKGCLKMNPDRDLTCSQLLESSYFDSFQEAQIKRK 329
Db 261 DLMLQLQYDPSQRITAKTALAHYPFSSPEPSPAARQ 297

RESULT 4
I50474
protein kinase (EC 2.7.1.37) cdc2 [similarity] - goldfish
C:Species: Carassius auratus (goldfish)
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 05-May-2000
C:Accession: I50474
R:Kajiura, H.; Yamashita, M.; Katsu, Y.; Nagahama, Y.
Dev. Growth Differ. 35, 647-654, 1993
A:Title: Isolation and characterization of goldfish cdc2, a catalytic component of ma
A:Reference number: I50474
A:Accession: I50474
A:Status: translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-302 <KAJ>
A:Cross-references: GB:D17758; NID:g471097; PIDN:BAA04605.1; PID:g471098
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphotransferase
F:2-256/Domain: protein kinase homology <KIN>
F:10-18/Region: protein kinase ATP-binding motif
F:33,51,128,130/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 30.9%; Score 563; DB 1; Length 302;
Best Local Similarity 41.2%; Pred. No. 1.9e-19;
Matches 117; Conservative 60; Mismatches 71; Indels 36; Gaps 6;

QY 1 MEKYEKLAKTGEISYGVVFKCRNKTSGQVAVKKF-VESEDDPVVKKIALREIRMLKQLK 59
Db 1 MDDYLKIEKIGEGTYGVVYGRNKTGTGQVAMKKIPLSEEEG-VPSTAVREISLLKELQ 59

QY 60 HPNLVNLIEVFRKRKMHVFEYCDHTLLNELEPNMG--VADGVTKSVLWQTLQALNFC 117
Db 60 HPNVVRLLDVLMQESKLYLVEFLSMDLKKYLDSPSGQFMDPMLVKSYLEIGILFC 119

QY 118 HIHNCIHRDIKIPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPPELLVGD 176
Db 120 HCRVLHRDLKPNQLLIDNKGVIKADFGLARAFGVPRVYTHEVVTLWYRAPEVLLGAS 179

QY 177 QYGSVDIWAIGCVFAELLTGQPLWPKSDVDQLYLIIRTL----- 217
Db 180 RYSTPVDVWSIGTIFAELATKKPLFHGDSIDQLFRIFRTLTGTPNNEVWPDVESLPDYKN 239

QY 233 LTSSDPPAVASQSAGITGKLIPRHQSIFKSNQFFHGIS-IP-----EPEDMETLEE 282
Db 222 -----TPNEQS-----WPGVSCLPDFKTAFPWQAQDLATV-- 252
QY 283 KFSVDHPVALNFMKGOLKMNPDRLTCSQLLESSYFDSFQEAQ 325
Db 253 -VPNLDPAGLDLLSKMLRYEPSKRITARQALEHEYFKDLEVQ 294

RESULT 12
T49271
CELL DIVISION CONTROL PROTEIN 2 HOMOLOG A - Arabidopsis thaliana
N;Alternate names: protein T21J18.20
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jun-2000
C;Accession: T49271
R;Rieger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Mewes, H.W.; Rudd, S.
submitted to the Protein Sequence Database, April 2000
A;Reference number: 225021
A;Accession: T49271
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-294 <RIE>
A;Cross-references: EMBL:AL132963; GSPDB:GN00061; ATSP:T21J18.20
A;Experimental source: cultivar Columbia; BAC clone T21J18
C;Genetics:
A;Gene: ATSP:T21J18.20
A;Map position: 3
A;Introns: 3/3; 67/3; 105/3; 163/3; 218/2; 247/3; 265/3
C;Superfamily: kinase-related transforming protein; protein kinase homology

Query Match 30.0%; Score 546.5; DB 2; Length 294;
Best Local Similarity 35.2%; Pred. No. 1.1e-18;
Matches 114; Conservative 71; Mismatches 96; Indels 43; Gaps 8;

QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MDQYEKVEKIGEGTYGVVYKARDKVTNETIALKKIRLEQDEGVPSTAIRESILKEMQH 60
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELERNPNQVAD-GVIKSVLWQTLQALNFCHI 119
Db 61 SNIVKYDDVVHSEKRLYLVEYLDLCLKKMDSTPDFSKDLHMIKTYLYQILRGTA YCHS 120
QY 120 HNCIHRDIKPENILITKQ-GIIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDQ 177
Db 121 HRVLHRLDKPQNLLIDRRNTSLKLADFGLARAFGIPVRTFTHEVVTLWYRAPELLGSHH 180
QY 178 YGSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIR---TLVETGFRHVDQAGLELLT 234
Db 181 YSTPVDIWSVGCIFAEMISQKPLFPDGDSEIDQLFKIPRMGTPYEDTWR----- 229
QY 235 SSDPPAVASQSAGITGKLIPRHQSIFKSNQFFHGISIPEDMETLEEKFSVDHPVALNF 294
Db 230 -----GVTS--LPDYKSAPPK-----WKPTDLETF---VPNLDPDGGVDL 263
QY 295 MKGCLKMNPDDRLTCSQLLESSYF 318
Db 264 LSKMLLMDPTKRINARAALHEYF 287

RESULT 13
A29539
protein kinase (EC 2.7.1.37) cdc2 - human
N;Alternate names: cell division control protein 2 (CDC2)
C;Species: Homo sapiens (man)
C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 21-Jul-2000
C;Accession: A29539
R;Lee, M.G.; Nurse, P.
Nature 327, 31-35, 1987
A;Title: Complementation used to clone a human homologue of the fission yeast cell cycle
A;Reference number: A29539; MUID:87201915

A;Accession: A29539
A;Molecule type: mRNA
A;Residues: 1-297 <LEE>
A;Cross-references: GB:X05360; NID:g29838; PIDN:CAA28963.1; PID:g29839
C;Genetics:
A;Gene: GDB:CDC2
A;Cross-references: GDB:119052; OMIM:116940
A;Map position: 10q21.1-10q21.1
C;Superfamily: kinase-related transforming protein; protein kinase homology
F;2-256/Domain: protein kinase homology <KIN>
F;10-18/Region: protein kinase ATP-binding motif
F;14,161/Binding site: phosphate (Thr) (covalent) #status predicted
F;15/Binding site: phosphate (Tyr) (covalent) #status predicted
F;33,51,128,130/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 29.9%; Score 545; DB 2; Length 297;
Best Local Similarity 40.2%; Pred. No. 1.3e-18;
Matches 117; Conservative 58; Mismatches 74; Indels 42; Gaps 7;

QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVVAVKKF-VESEDDPVVKKIALREIRMLKQLK 59
Db 1 MEDYTKIEKIGEGTYGVVYKGRHKTGTQGVVAMKKIRLESEEG-VpSTAIRESILKELR 59
QY 60 HPNLVNLIEVFRKRKMHVFEYCDHTLLNELERNPNG--VADGVIKSVLWQTLQALNFC 117
Db 60 HPNIVSLQDVLMDQSRLYLIFEFLSMDLKKYLDSPPGQYMDSSLVKSYYLQILQGIVFC 119
QY 118 HINCIIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGD 176
Db 120 HSRVLHRLDKPQNLLIDDKGTIKLADFGLARAFGIPRVYTHEVVTLWYRSPEVLLGSA 179
QY 177 QYSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTL----- 217
Db 180 RYSTPVDIWSIGTIFAELATKKPLFHGDSEIDQLFRIFRALGTPNNEVWPEVESLQDYKN 239
QY 218 -----VETGFRHVDQAGLELLTSS---DPPAVASQSAGITGKLIPRH 256
Db 240 TFPKWPGSLASHVKNLDENGLDLLSKMLIYDP-----AKRISGKMALNH 284

RESULT 14
A44878
protein kinase (EC 2.7.1.37) cdk2 [validated] - goldfish
C;Species: Carassius auratus (goldfish)
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 05-May-2000
C;Accession: A44878
R;Hirai, T.; Yamashita, M.; Yoshikuni, M.; Tokumoto, T.; Kajiura, H.; Sakai, N.; Naga
Dev. Biol. 152, 113-120, 1992
A;Title: Isolation and characterization of goldfish cdk2, a cognate variant of the ce
A;Reference number: A44878; MUID:92331802
A;Accession: A44878
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-298 <HIR>
A;Cross-references: GB:S40289; NID:g251619; PIDN:AAB22550.1; PID:g251620
A;Experimental source: oocyte
A;Note: sequence extracted from NCBI backbone (NCBIN:108782, NCBIP:108783)
C;Genetics:
A;Gene: cdk2
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; cell cycle control; mitosis; phosphoprotein; phosphotransferase; ser
F;2-255/Domain: protein kinase homology <KIN>
F;10-18/Region: protein kinase ATP-binding motif
F;33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 29.9%; Score 545; DB 1; Length 298;
Best Local Similarity 46.6%; Pred. No. 1.3e-18;
Matches 102; Conservative 48; Mismatches 67; Indels 2; Gaps 2;

QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60

Db 1 MESFQVEKIGEGTYGVYKAKNKVTGETVALKKIRLDTETEGVPSTAIREISLLKELNH 60
Qy 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEARNP-NGVADGVVKSVLWQTLQALNFCHI 119
Db 61 PNIIVLHDVHTENKLYLVFEFLHQDLKRFMDSSVTGISLPLVKSILFQLLQGLAFCHS 120
Qy 120 HNCIHRDIKPENILITKQGIKICDFGFAQL-IPGDYTDYVATRWYRAPELLVGDTQY 178
Db 121 HRVLHRDLKPQNLLINAQGEIKLADFGLARAFGVPVRVYTHEVVTWYRAPELLGCKYY 180
Qy 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTL 217
Db 181 STAVDIWSLGCIFAEMITRKALFPGDSEIDQLFRIFRTL 219

RESULT 15
I45977
cyclin-dependent kinase 1 - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 19-Dec-1997 #sequence_revision 19-Dec-1997 #text_change 18-Jun-1999
C:Accession: I45977
R:Yang, L.; Farin, C.E.
Gene 141, 283-286, 1994
A:Title: Identification of cDNAs encoding bovine cyclin B and Cdk1/cdc2.
A:Reference number: I45977; MUID:94215918
A:Accession: I45977
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-297 <YAN>
A:Cross-references: GB:L26547; NID:g433155; PIDN:AAAL8894.1; PID:g498173
C:Genetics:
A:Gene: cdk1/cdc2
C:Superfamily: kinase-related transforming protein; protein kinase homology
F;2-256/Domain: protein kinase homology <KIN>

Query Match 29.8%; Score 542; DB 2; Length 297;
Best Local Similarity 40.2%; Pred. No. 1.7e-18;
Matches 117; Conservative 57; Mismatches 75; Indels 42; Gaps 7;
Qy 1 MEKYELAKTGESYGVVFCRNKTSQGVVAVKKF-VESEDDPVVKKIALREIRMLKQLK 59
Db 1 MEDYTKIEKIGEGTYGVYKGRHKTGQVVAMKKIRLESEEG-VPSTAIREISLLKELR 59
Qy 60 HPNLVNLEVFRRKRKMHVFEYCDHTLLNELEARNP--VADGVVKSVLWQTLQALNFC 117
Db 60 HPNIVSLQDVLMDQSRLYLIFEFLSMDLKKYLDSSIPGQFMDSLSLVKSILYQILQGIVFC 119
Qy 118 HHCNCHRIDIKPENILITKQGIKICDFGFAQL-IPGDYTDYVATRWYRAPELLVGDT 176
Db 120 HSRRVLHRDLKPQNLLIDDKGTIKLADFGLARAFGPIRVYTHEVVTWYRSPEVLLGSA 179
Qy 177 QYGSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTL----- 217
Db 180 GYSTPVDIWSIGTIFAELATKKPLFHGDSEIDQLFRIFRALGTPNNEVWPEVESLQDYKS 239
Qy 218 -----VETGFRHVDOAGLELLTSS---DPPAVASQSAGITGKLIPIRH 256
Db 240 TFPKWKPGSLASHVKNLDENGLDLSKMLIYDP-----AKRISGKMALNH 284

Search completed: May 3, 2002, 11:03:16
Job time: 110 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2002, 11:02:46 ; Search time 17.01 Seconds
(without alignments)
747.954 Million cell updates/sec

Title: US-09-671-050-10
Perfect score: 1820
Sequence: 1 MEYKELAKTGEYSYGVVFK.....RKARNEGRNRRRQQVLPLKS 347

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1157.5	63.6	358	1	KKIA_HUMAN	Q00532 homo sapien
2	694.5	38.2	1030	1	STK9_HUMAN	O76039 homo sapien
3	565	31.0	305	1	CDK3_HUMAN	Q00526 homo sapien
4	563	30.9	302	1	CC2_CARAU	P51958 carassius a
5	561	30.8	297	1	CDK2_XENLA	P23437 xenopus lae
6	558.5	30.7	292	1	CC2H_DICDI	P34117 dictyosteli
7	554	30.4	298	1	CDK2_HUMAN	P24941 homo sapien
8	551.5	30.3	294	1	CC2I_ORYSA	P29618 oryza sativ
9	551.5	30.3	294	1	CC2A_ARATH	P24100 arabidopsis
10	550	30.2	298	1	CDK2_CRIGR	O55076 cricetulus
11	548	30.1	298	1	CDK2_RAT	Q63699 rattus norv
12	546	30.0	298	1	CDK2_MESAU	P48963 mesocricetu
13	545	29.9	297	1	CC2_HUMAN	P06493 homo sapien
14	545	29.9	298	1	CDK2_CARAU	P43450 carassius a
15	542	29.8	297	1	CC2_BOVIN	P48734 bos taurus
16	541.5	29.8	294	1	CC2_MAIZE	P23111 zea mays (m
17	541	29.7	297	1	CC2_RAT	P39951 rattus norv
18	540.5	29.7	297	1	CC2_DROME	P23572 drosophila
19	540.5	29.7	302	1	CC22_XENLA	P24033 xenopus lae
20	540	29.7	302	1	CC21_XENLA	P35567 xenopus lae
21	539	29.6	303	1	CC2_CHICK	P13863 gallus gall
22	537	29.5	288	1	CC2H_PLAFK	Q07785 plasmodium
23	536.5	29.5	294	1	CC2_VIGUN	P52389 vigna ungui
24	536.5	29.5	311	1	CC23_TRYBB	P54666 trypanosoma
25	532	29.2	297	1	CC2_MOUSE	P11440 mus musculu
26	532	29.2	301	1	CC21_TRYBB	P38973 trypanosoma
27	531	29.2	292	1	CDK5_BOVIN	Q02399 bos taurus
28	531	29.2	292	1	CDK5_RAT	Q03114 rattus norv
29	530.5	29.1	294	1	CC22_MEDSA	Q05006 medicago sa
30	529	29.1	292	1	CDK5_MOUSE	P49615 mus musculu
31	527	29.0	301	1	CC21_TRYCO	P54664 trypanosoma
32	526	28.9	292	1	CDK5_HUMAN	Q00535 homo sapien
33	526	28.9	296	1	CC2_DICDI	P34112 dictyosteli

34	525	28.8	294	1	CDK5_DROME	P48609 drosophila
35	521.5	28.7	294	1	CC2A_ANTMA	Q38772 antirrhinum
36	521	28.6	292	1	CDK5_XENLA	P51166 xenopus lae
37	519.5	28.5	346	1	CDK7_HUMAN	P50613 homo sapien
38	518	28.5	292	1	CC22_ORYSA	P29619 oryza sativ
39	518	28.5	305	1	PH85_YEAST	P17157 saccharomyc
40	517.5	28.4	346	1	CDK7_MOUSE	Q03147 mus musculu
41	516.5	28.4	314	1	CC2C_DROME	P23573 drosophila
42	516	28.4	346	1	CDK2_MOUSE	P97377 mus musculu
43	512	28.1	304	1	PH85_KLULA	Q92241 kluyveromyc
44	510	28.0	332	1	CC2_CAEEL	P34556 caenorhabdi
45	510	28.0	345	1	CC22_TRYBB	P54665 trypanosoma

ALIGNMENTS

RESULT 1

ID	KKIA_HUMAN	STANDARD;	PRT;	358 AA.
AC	Q00532;			
DT	01-APR-1993 (Rel. 25, Created)			
DT	01-APR-1993 (Rel. 25, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	SERINE/THREONINE-PROTEIN KINASE KKIALRE (EC 2.7.1.1-) (CYCLIN-DEPENDENT KINASE-LIKE 1).			
GN	CDKL1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92347325; PubMed=1639063;			
RA	Meyerson M., Enders G.H., Wu C.-L., Su L.-K., Gorka C., Nelson C., Harlow E., Tsai L.-H.;			
RT	"A family of human cdc2-related protein kinases.";			
RL	EMBO J. 11:2909-2917(1992).			
CC	-!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.			
CC	CDC2/CDKX SUBFAMILY.			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC	EMBL; X66358; CAA47002.1; "	
DR	EMBL; X66359; CAA47002.1; JOINED.	
DR	PIR; S22744; S22744.	
DR	PIR; S22745; S22745.	
DR	PIR; S23383; S23383.	
DR	HSSP; P24941; IAQ1.	
DR	MIM; 603441; "	
DR	InterPro; IPR000719; Euk_pkinase.	
DR	InterPro; IPR002290; Ser_thr_kin_actsite.	
DR	Pfam; PF00069; pkinase; 1.	
DR	SMART; SM00220; S_TKC; 1.	
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.	
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.	
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.	
KW	Transferase; Serine/threonine-protein kinase; ATP-binding.	
FT	DOMAIN 5 288	
FT	NP_BIND 11 19	
FT	BINDING 34 34	
FT	ACT_SITE 127 127	
FT	SEQUENCE 358 AA; 88344321P24B77C6 CRC64;	

Query Match 63.6%; Score 1157.5; DB 1; Length 358;
Best Local Similarity 64.7%; Pred. No. 1.3e-71;

QY 60 HPNLVNLIEVFRKRKMHVFEYCDHTLLNELRPNNG--VADGVIKSVLWQTLQALNFC 117
Db 60 HPNVVRLLDVLMQESKLYLVFEFLSMDLKKYLDSPSGQFMDPMLVKSVLYQILEGILFC 119
QY 118 HIHNCIHRDIKPNILITKQGIKICDFGFAQL-IPGDAYTDYVATRWYRAPELLVGD 176
Db 120 HCRVLRDLKPQNLLIDNKGVIKLAFLADGFLARAFGVPRVYTHEVVTWLYRAPEVLGAS 179
QY 177 QYGSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTL----- 217
Db 180 RYSTPVDVWSIGTIFAELATKKPLFHGDSEIDQLFRIFRITLGTNPNEVWPDVESLPDYKN 239
QY 218 -----VETGFRHVDQAGLELLTSS---DPPAVASQSAGIT 249
Db 240 TFPKWKSGNLAFTVKNLDKNGIDLLTKMLIYDPPKRISARQAMT 283

RESULT 5
CDK2_XENLA STANDARD; PRT; 297 AA.
AC P23437;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CELL DIVISION PROTEIN KINASE 2 (EC 2.7.1.-) (CDC2 HOMOLOG EG1 PROTEIN KINASE).
GN EG1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=91126051; PubMed=1704128;
RA Paris J., le Guellec R., Couturier A., le Guellec K., Omilli F.,
RA Camonis J., Macneill S., Philippe M.;
RT "Cloning by differential screening of a xenopus cDNA coding for a
RT protein highly homologous to cdc2.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:1039-1043(1991).
RN [2]
RP PHOSPHORYLATION AT THR-160.
RX MEDLINE=93345457; PubMed=8393783;
RA Poon R.Y.C., Yamashita K., Adamczewski J.P., Hunt T., Shuttleworth J.;
RA "The cdc2-related protein p40MO15 is the catalytic subunit of a
RT protein kinase that can activate p33cdc2 and p34cdc2.";
RL EMBO J. 12:3123-3132(1993).
CC -!- FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.
CC INTERACTS WITH CYCLINS A, D, OR E.
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-160 ACTIVATES IT (BY
CC SIMILARITY).
CC -!- DEVELOPMENTAL STAGE: SYNTHESIZED IN UNFERTILIZED EGG, BUT NO
CC LONGER MADE IN THE EARLY EMBRYO.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X14227; CAA32443.1; -.
DR PIR; A37871; A37871.
DR HSSP; P24941; IHCK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_kin_actsite.
DR Pfam; PF000069; pkinase; 1.
DR SMART; SM00220; S_TKc; 1.

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW transferase; Serine/threonine-protein kinase; ATP-binding;
KW Cell cycle; Cell division; Mitosis; Phosphorylation.
FT DOMAIN 4 286 PROTEIN KINASE.
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 127 127 BY SIMILARITY.
FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 160 160 PHOSPHORYLATION (BY CAK).
SQ SEQUENCE 297 AA; 33969 MW; EC30204FCB8D198C CRC64;

Query Match 30.8%; Score 561; DB 1; Length 297;
Best Local Similarity 38.2%; Pred. No. 2.4e-31;
Matches 116; Conservative 59; Mismatches 89; Indels 40; Gaps 6;
QY 1 MEKYEKLAKTGESYGVVFKCRNKTSQVVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MENFOKVEKIGEGTYGVVYKARNRETGEIVALKKIRLDTETEGVPSTAIRSLLKELNH 60
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELER-NPNGVADGVIKSVLWQTLQALNFC 119
Db 61 PNIVKLLDVHTENKLYLVFEFLNQDLKKFMDRSNISGLALVKSYLFQLLQGLAFCHS 120
QY 120 HNCIHRDIKPNILITKQGIKICDFGFAQL-IPGDAYTDYVATRWYRAPELLVGD 178
Db 121 HRVLRDLKPQNLLINS DGAIKLADFLARAFGVPRVYTHEVVTWLYRAPELLGCKFY 180
QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTL----- 217
Db 181 STAVDIWSLGCIFAEMITRRALFPDSEIDQLFRIFRITLGTDPVSWPGVTTMPDYKSTF 240
QY 218 ---VETGFRHV---DQAGLELLTSSDPPAVASQSAGITKLI PRHQSIKSN GFFHGIS 270
Db 241 PKWIRQDFSKVVPPLDE DGRLLAQM---LQYDSNKRISAKVALTHP-----FFRDVS 290
QY 271 IPEP 274
Db 291 RPTP 294

RESULT 6
CC2H_DICDI
ID CC2H_DICDI STANDARD; PRT; 292 AA.
AC P34117;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CDC2-LIKE SERINE/THREONINE-PROTEIN KINASE CRP (EC 2.7.1.-).
GN CRPA OR CRP.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94032415; PubMed=8218353;
RA Michaelis C.E., Weeks G.;
RT "The isolation from a unicellular organism, Dictyostelium discoideum,
RT of a highly-related cdc2 gene with characteristics of the PCTAIRE
RT subfamily.";
RL Biochim. Biophys. Acta 1179:117-124(1993).
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; L00652; AAAL6056.1; -.
DR PIR; S40021; S40021.
DR HSP; P24941; 1AQ1.
DR DictyDb; DD05039; crpA.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_kin_actsite.
DR Pfam; PF00069; pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW transferase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation.
FT DOMAIN 4 285 PROTEIN KINASE.
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 126 126 BY SIMILARITY.
FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 292 AA; 33226 MW; 361AB54C4E3BD41E CRC64;

Query Match 30.7%; Score 558.5; DB 1; Length 292;
Best Local Similarity 35.0%; Pred. No. 3.4e-31;
Matches 115; Conservative 66; Mismatches 101; Indels 47; Gaps 6;

QY 1 MEKYEKLAKTGE GSYGVVFKCRNKTSGQWAVKFKVESEDDPVVKKIALFREIMLKQLKH 60
Db 1 MEKYSKIEKLGE GTYGIVNKAKNRETGEIVALKRIRLDSEDEGVPCPTAIREISLLKELKH 60
QY 61 PNLVNLIEVFRKRKMHLYFVEYCDHTLLNELERNPNGVADGVIKSVLWQTLQALNFCUIH 120
Db 61 PNIVRLHDV IHTERKLTLVPEYLDQDLKKYLDCEGGEISKPTIKSFMYQLLKGVAFCDDH 120
QY 121 NCIHRIKPKENILITTKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPPELLVGDTOYG 179
Db 121 RVLHRDLKPQNLLINRKGEIKLADFLGARAFAFGIPVRTYSHEVTVLWYRAPDVLMGSRKYS 180
QY 180 SSVDTWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLVETGFRIVDQAGLELLTSSDPP 239
Db 181 TPIDIWSALCIFAEMASGRPLFGSGTSDQLFRIFKILGTP-----NEESWP 227
QY 240 AVASQSAGITGKLIPRHQSIF-----KSNFFHGISIPEPEDMETLEEKFSVDHPVALN 293
Db 228 SITE-----LPEYKTFDPVHPAHQLSSIVHG-----LDEK-----GLN 260
QY 294 FMKGCLKNPDDRLTCSQLLESYFDSFQ 322
Db 261 LLSKMLQYDPNQRIITAAALKHPYFDGLE 289

RESULT 7
CDK2_HUMAN
ID CDK2_HUMAN STANDARD; PRT; 298 AA.
AC P24941;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CELL DIVISION PROTEIN KINASE 2 (EC 2.7.1.-) (P33 PROTEIN KINASE).
GN CDK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91330891; PubMed=1714386;
RA Elledge S.J., Spottswood M.R.;
RT "A new human p34 protein kinase, CDK2, identified by complementation

RT of a cdc28 mutation in Saccharomyces cerevisiae, is a homolog of
RT Xenopus Egl.";
RL EMBO J. 10:2653-2659(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91367262; PubMed=1653904;
RA Tsai L.-H., Harlow E., Meyerson M.;
RT "Isolation of the human cdk2 gene that encodes the cyclin A- and
RT adenovirus E1A-associated p33 kinase.";
RL Nature 353:174-177(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92020980; PubMed=1717994;
RA Ninomiya-Tsuji J., Nomoto S., Yasuda H., Reed S.I., Matsumoto K.;
RT "Cloning of a human cDNA encoding a CDC2-related kinase by
RT complementation of a budding yeast cdc28 mutation.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:9006-9010(1991).
RN [4]
RP PHOSPHORYLATION SITES.
RX MEDLINE=93010995; PubMed=1396589;
RA Gu Y., Rosenblatt J., O'Morgan D.O.;
RT "Cell cycle regulation of CDK2 activity by phosphorylation of Thr160
RT and Tyr15.";
RL EMBO J. 11:3995-4005(1992).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RX MEDLINE=93288132; PubMed=8510751;
RA de Bondt H.L., Rosenblatt J., Jancarik J., Jones H.D.,
RA Morgan D.O., Kim S.-H.;
RT "Crystal structure of cyclin-dependent kinase 2.";
RL Nature 363:595-602(1993).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF COMPLEX WITH CYCLIN A.
RX MEDLINE=95356811; PubMed=7630397;
RA Jeffrey P.D., Russo A.A., Polyak K., Gibbs E., Hurwitz J.,
RA Massague J., Pavletich N.P.;
RT "Mechanism of CDK activation revealed by the structure of a
RT cyclinA-CDK2 complex.";
RL Nature 376:313-320(1995).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.33 ANGSTROMS) OF COMPLEX WITH L868276.
RX MEDLINE=96181476; PubMed=8610110;
RA de Azevedo W.F. Jr., Muleer-Dieckmann H.-J., Schulze-Gahmen U.,
RA Worland P.J., Sausville E., Kim S.-H.;
RT "Structural basis for specificity and potency of a flavonoid
RT inhibitor of human CDK2, a cell cycle kinase.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:2735-2740(1996).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF COMPLEX WITH KIP1.
RX MEDLINE=96300318; PubMed=8684460;
RA Russo A.A., Jeffrey P.D., Patten A.K., Massague J., Pavletich N.P.;
RT "Crystal structure of the p27kip1 cyclin-dependent-kinase inhibitor
RT bound to the cyclin A-Cdk2 complex.";
RL Nature 382:325-331(1996).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF COMPLEX WITH CG2A.
RX MEDLINE=96313126; PubMed=8756328;
RA Russo A.A., Jeffrey P.D., Pavletich N.P.;
RT "Structural basis of cyclin-dependent kinase activation by
RT phosphorylation.";
RL Nat. Struct. Biol. 3:696-700(1996).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=97075215; PubMed=8917641;
RA Schulze-Gahmen U., de Bondt H.L., Kim S.-H.;
RT "High-resolution crystal structures of human cyclin-dependent kinase
RT 2 with and without ATP: bound waters and natural ligand as guides for
RT inhibitor design.";
RL J. Med. Chem. 39:4540-4546(1996).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=97475219; PubMed=9334743;
RA Lawrie A.M., Noble M.E.M., Tunnah P., Brown N.R., Johnson L.N.,

RA Endicott J.A.;

RT "protein kinase inhibition by staurosporine revealed in details of

RL the molecular interaction with CDK2.";

RL Nat. Struct. Biol. 4:796-801(1997).

RN [12]

RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF COMPLEX WITG CKS1.

RX MEDLINE=96182647; PubMed=8601310;

RA Bourne Y., Watson M.H., Hickey M.J., Holmes W., Rocque W., Reed S.I.,

RA Tainer J.A.;

RT "Crystal structure and mutational analysis of the human CDK2 kinase

RT complex with cell cycle-regulatory protein CksHs1.";

RL Cell 84:863-874(1996).

RN [13]

RP X-RAY CRYSTALLOGRAPHY (2.05 ANGSTROMS).

RX MEDLINE=98342369; PubMed=9677190;

RA Gray N.S., Wodicka L., Thunnissen A.-M.W.H., Norman T.C., Kwon S.,

RA Espinoza F.H., Morgan D.O., Barnes G., Leclerc S., Meijer L.,

RA Kim S.H., Lockhart D.J., Schultz P.G.;

RT "Exploiting chemical libraries, structure, and genomics in the search

RT for kinase inhibitors.";

RL Science 281:533-538(1998).

CC -!- FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.

CC INTERACTS WITH CYCLINS A, D, OR E. ACTIVITY OF CDK2 IS MAXIMAL

CC DURING S PHASE AND G2.

CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES

CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-160 ACTIVATES IT.

CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

CC CDC2/CDKX SUBFAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X61622; CAA43807.1; -.

DR EMBL; X62071; CAA43985.1; -.

DR EMBL; M68520; AAA35667.1; -.

DR PIR; A41227; A41227.

DR PIR; S16520; S16520.

DR PIR; S17873; S17873.

DR PDB; 1FIN; 27-JAN-97.

DR PDB; 1HCK; 07-DEC-96.

DR PDB; 1HCL; 07-DEC-96.

DR PDB; 1AQL; 12-NOV-97.

DR PDB; 1JST; 11-JAN-97.

DR PDB; 1JSU; 29-JUL-97.

DR PDB; 1BUH; 09-SEP-98.

DR PDB; 1B38; 23-DEC-98.

DR PDB; 1B39; 23-DEC-98.

DR PDB; 1CKP; 13-JAN-99.

DR MIM; 116953; -.

DR InterPro; IPR000719; Euk_pkinase.

DR InterPro; IPR002290; Ser_thr_kin_actsite.

DR Pfam; PF00069; pkinase; 1.

DR SMART; SM00220; S_TKc; 1.

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.

DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.

KW Transferase; Serine/threonine-protein kinase; ATP-binding;

KW Cell cycle; Cell division; Mitosis; Phosphorylation; 3D-structure.

FT DOMAIN 4 286

FT NP_BIND 10 18

FT BINDING 33 33

FT ACT_SITE 127 127

FT MOD_RES 14 14

FT MOD_RES 15 15

FT MOD_RES 160 160

FT MUTAGEN 14 14

FT MUTAGEN 15 15

FT MUTAGEN 160 160

FT T->A: INCREASE ACTIVITY 2 FOLD.

FT Y->F: INCREASE ACTIVITY 2 FOLD.

FT T->A: ABOLISHES ACTIVITY.

SQ SEQUENCE 298 AA; 33929 MW; F90A0F4E70910B51 CRC64;

Query Match 30.4%; Score 554; DB 1; Length 298;

Best Local Similarity 38.4%; Pred. NO. 7.1e-31;

Matches 118; Conservative 54; Mismatches 89; Indels 46; Gaps 6;

QY 1 MEKYEKLAKTGE GSYGVVFKCRNKTSGOVAVKKEVSEDDPVVKKIALREIRMLKQLKH 60

Db 1 MENFQKVEKIGEGTYGVVYKARNKLTGEVVALKKIRLDTETEGVVPSTAIRISLKLKLNH 60

QY 61 PNLVNLIEVFRKRKRMILVFECYCDHTLLNELERNP-NGVADGVIKSVLWQTLQALNFCHI 119

Db 61 PNIVKLLDVHTENKLYLVFEFLHQDLKKEMDASALTGIPPLIKSYLFQLLOGLAFCHS 120

QY 120 HNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDATDYVATRWYRAPELLVGDYQY 178

Db 121 HRVLHRDLKPQNLLINTEGAIKLADFLARAFGVPRVRYTHEVVTLMYRAPEILLGCKYY 180

QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTL----- 217

Db 181 STAVDIWSLGCIFAEMVTRRALFPDSEIDQLFRIFRTLTGTPDEVVWPGVTSMPDYKPSF 240

QY 218 ---VETGFRHV----DQAGLELLTSS---DPPAVASQSAGITGKLIPRHQSIEKSNGFHH 267

Db 241 PKWARQDFSKVVPPLDEDEGRKSLLSQMLHYDPNKRISAKAAL-----AHPFFQ 287

QY 268 GISIPEP 274

Db 288 DVTKPVP 294

RESULT 8

CC21_ORYSA

ID CC21_ORYSA STANDARD; PRT; 294 AA.

AC P29618;

DT 01-APR-1993 (Rel. 25, Created)

DT 01-APR-1993 (Rel. 25, Last sequence update)

DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE CELL DIVISION CONTROL PROTEIN 2 HOMOLOG 1 (EC 2.7.1.-).

GN CDC2-1.

OS Oryza sativa (Rice).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartioideae; Oryzeae; Oryza.

OX NCBI_TaxID=4530;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. NIPPONBARE;

RX MEDLINE=92293101; PubMed=1376401;

RA Hashimoto J., Hirabayashi T., Hayano Y., Hata S., Ohashi Y.,

RA Suzuki I., Utsugi T., Toh-E A., Kikuchi Y.;

RT "Isolation and characterization of cDNA clones encoding cdc2

RT homologues from Oryza sativa: a functional homologue and cognate

RT variants.";

RL Mol. Gen. Genet. 233:10-16(1992).

CC -!- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC

CC CELL CYCLE. COMPONENT OF THE KINASE COMPLEX THAT PHOSPHORYLATES

CC THE REPETITIVE CARBOXYL-TERMINUS OF RNA POLYMERASE II.

CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES

CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-161 ACTIVATES IT (BY

CC SIMILARITY).

CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

CC CDC2/CDKX SUBFAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----


```
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D28753; BAA05947.1; -.
CC EMBL; D63162; BAA09638.1; -.
CC InterPro; IPR000719; Euk_pkinase.
CC InterPro; IPR002290; Ser_thr_kin_actsite.
CC Pfam; PF00069; pkinase; 1.
CC SMART; SM00220; S_TKC; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC Transferase; Serine/threonine-protein kinase; ATP-binding; Cell cycle;
CC Cell division; Mitosis; Phosphorylation; Alternative splicing.
CC DOMAIN 4 286 PROTEIN KINASE.
CC NP_BIND 10 18 ATP (BY SIMILARITY).
CC BINDING 33 33 ATP (BY SIMILARITY).
CC ACT_SITE 127 127 BY SIMILARITY.
CC MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
CC MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
CC MOD_RES 160 160 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).
CC CONFLICT 79 79 V -> C (IN REF. 2).
CC CONFLICT 99 99 L -> I (IN REF. 2).
CC CONFLICT 124 124 L -> C (IN REF. 2).
CC SEQUENCE 298 AA; 33887 MW; C8CB3ADCE9B97F88 CRC64;

Query Match 30.1%; Score 548; DB 1; Length 298;
Best Local Similarity 38.1%; Pred. No. 1.8e-30;
Matches 117; Conservative 55; Mismatches 89; Indels 46; Gaps 6;

QY 1 MEKYEKLAKTGEQSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MENFQKVEKIGEGTYGVVYKAKNKLGTGEVVALKIRLDTETEGVPSTAIRESILLKELNH 60

QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEARNP-NGVADGVIKSVLWQTLQALNFCHI 119
Db 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEARNP-NGVADGVIKSVLWQTLQALNFCHI 119

QY 120 HNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDQY 178
Db 121 HRVLHRLDKPQNLLINAECSIKLADFGLARAFGVPRVTYTHEVVTLWYRAPEILLGCKYY 180

QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTL----- 217
Db 181 STAVDIWSLGCIFAEMVTRRALFPGDSEIDQLFRIFRTLGTDPDEVVWPGVTSMPDYKPSF 240

QY 218 ---VETGFRHV---DQAGLELLTSS---DPPAVASQSAGITGKLIIPRHQSIFKSNNGFFH 267
Db 241 PKWARQDFSKVVPPLDEDEGRSLLSQMLHYDPNKRISAKAAL-----AHPFFQ 287

QY 268 GISIPEP 274
Db 288 DVTKPVP 294

RESULT 12
CDK2_MESAU STANDARD; PRT; 298 AA.
ID CDK2_MESAU
AC P48963;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CELL DIVISION PROTEIN KINASE 2 (EC 2.7.1.-).
GN CDK2.
```

```
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
MEDLINE=94107363; PubMed=8280171;
RA Noguchi E., Sekiguchi T., Yamashita K., Nishimoto T.;
RT "Molecular cloning and identification of two types of hamster cyclin-
RT dependent kinases: cdk2 and cdk2L.";
RL Biochem. Biophys. Res. Commun. 197;1524-1529(1993).
CC -!- FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.
CC INTERACTS WITH CYCLINS A, D, OR E. ACTIVITY OF CDK2 IS MAXIMAL
CC DURING S PHASE AND G2.
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-160 ACTIVATES IT (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D17350; BAA04165.1; -.
CC HSSP; P24941; 1AQL.
CC InterPro; IPR000719; Euk_pkinase.
CC InterPro; IPR002290; Ser_thr_kin_actsite.
CC Pfam; PF00069; pkinase; 1.
CC SMART; SM00220; S_TKC; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC Transferase; Serine/threonine-protein kinase; ATP-binding;
CC Cell cycle; Cell division; Mitosis; Phosphorylation.
KW DOMAIN 4 286 PROTEIN KINASE.
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 127 127 BY SIMILARITY.
FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 160 160 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).
FT SEQUENCE 298 AA; 33839 MW; D1LC036664C4C7CE CRC64;
SQ

Query Match 30.0%; Score 546; DB 1; Length 298;
Best Local Similarity 38.1%; Pred. No. 2.5e-30;
Matches 117; Conservative 54; Mismatches 90; Indels 46; Gaps 6;

QY 1 MEKYEKLAKTGEQSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MENFQKVEKIGEGTYGVVYKAKNKLGTGEVVALKIRLDTETEGVPSTAIRESILLKELNH 60

QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEARNP-NGVADGVIKSVLWQTLQALNFCHI 119
Db 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEARNP-NGVADGVIKSVLWQTLQALNFCHI 119

QY 120 HNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDQY 178
Db 121 HRVLHRLDKPQNLLINAECSIKLADFGLARAFGVPRVTYTHEVVTLWYRAPEILLGCKYY 180

QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTL----- 217
Db 181 STAVDIWSLGCIFAEMVTRRALFPGDSEIDQLFRIFRTLGTDPDEVVWPGVTSMPDYKPSF 240

QY 218 ---VETGFRHV---DQAGLELLTSS---DPPAVASQSAGITGKLIIPRHQSIFKSNNGFFH 267
Db 241 PKWARQDFSKVVPPLDEDEGRSLLSQMLHYDPNKRISAKAAL-----AHPFFQ 287
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2002, 11:02:21 ; Search time 40.57 Seconds
(without alignments)
1251.084 Million cell updates/sec

Title: US-09-671-050-10
Perfect score: 1820
Sequence: 1 MEKYEKLAKTGEVGVVFK.....RKARNEGRNRRRQVLPKLS 347

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues 473505
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_17:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1001.5	55.0	353	5 Q9U2H1	Q9u2h1 caenorhabdi
2	1001.5	55.0	392	5 Q9VMN3	Q9vmn3 drosophila
3	941.5	51.7	566	6 Q9TTK0	Q9ttk0 oryctolagus
4	915.5	50.3	493	4 Q92772	Q92772 homo sapien
5	904.5	49.7	564	11 Q9QYI2	Q9qyi2 mus musculu
6	904.5	49.7	568	11 Q9QUK0	Q9quk0 mus musculu
7	902	49.6	329	11 Q9QYI1	Q9qyi1 mus musculu
8	758.5	41.7	455	4 Q9P114	Q9pli4 homo sapien
9	755	41.5	657	5 Q9BMG2	Q9bmg2 trypanosoma
10	743.5	40.9	1106	5 Q9GRT9	Q9grt9 leishmania
11	737.5	40.5	457	11 Q9JMO2	Q9jmo2 rattus norv
12	737.5	40.5	505	11 Q9JMO1	Q9jmo1 rattus norv
13	618	34.0	1104	13 Q9W6R6	Q9w6r6 fugu rubrip
14	615.5	33.8	997	4 Q9UJL6	Q9ujl6 homo sapien
15	565.5	31.1	294	10 Q9ZRI1	Q9zri1 triticum ae
16	551	30.3	288	5 Q96821	Q96821 plasmodium
17	551	30.3	288	5 Q9XZD6	Q9xzd6 plasmodium
18	549.5	30.2	303	13 Q9DGA2	Q9dga2 oryzias jav
19	546.5	30.0	294	10 Q9M307	Q9m307 arabidopsis

20	546.5	30.0	303	13 Q9DGA1	Q9dga1 oryzias jav
21	545.5	30.0	294	10 Q82666	Q82666 brassica na
22	543.5	29.9	303	13 Q9DGD3	Q9dgd3 oryzias lat
23	542.5	29.8	294	10 P93101	P93101 chenopodium
24	542.5	29.8	300	5 O15890	O15890 toxoplasma
25	541.5	29.8	297	5 Q9TX69	Q9tx69 drosophila
26	541.5	29.8	303	13 Q9DGA5	Q9dga5 oryzias cur
27	541.5	29.8	303	13 Q9DG98	Q9d998 oryzias luz
28	540.5	29.7	297	5 Q9TX74	Q9tx74 drosophila
29	539.5	29.6	294	10 Q40790	Q40790 pinus conto
30	539.5	29.6	297	5 Q9TX73	Q9tx73 drosophila
31	539.5	29.6	300	5 O44000	O44000 toxoplasma
32	539.5	29.6	300	5 Q17066	Q17066 asterina pe
33	539	29.6	298	5 Q27032	Q27032 theileria p
34	538.5	29.6	294	10 Q40789	Q40789 petroselinu
35	538.5	29.6	294	10 O82135	O82135 pisum sativ
36	538.5	29.6	300	3 O13379	O13379 pneumocysti
37	538.5	29.6	300	3 O13380	O13380 pneumocysti
38	537.5	29.5	297	5 Q9TX71	Q9tx71 drosophila
39	537	29.5	288	5 O96820	O96820 plasmodium
40	536.5	29.5	297	5 Q9TX68	Q9tx68 drosophila
41	535.5	29.4	294	10 P93556	P93556 sesbania ro
42	535.5	29.4	294	10 Q41639	Q41639 vigna aconi
43	535.5	29.4	294	10 Q43361	Q43361 picea abies
44	534.5	29.4	294	10 Q40484	Q40484 nicotiana t
45	534.5	29.4	294	10 Q9FUR4	Q9fur4 nicotiana t

ALIGNMENTS

RESULT 1
Q9U2H1 ID Q9U2H1 PRELIMINARY; PRT; 353 AA.
AC Q9U2H1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Y42A5A.4 PROTEIN.
GN Y42A5A.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Steward C.A.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology."
RL Science 282:2012-2018(1998).
CC -I- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AL032618; CAB63367.1; -.
DR HSSP; P24941; 1HCL.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_kin_actsite.
DR Pfam; PF00069; pkinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 353 AA; 40665 MW; 7E32260D9849CAC9 CRC64;

Query Match 55.0%; Score 1001.5; DB 5; Length 353;
Best Local Similarity 53.4%; Pred. NO. 4.6e-80;
Matches 187; Conservative 59; Mismatches 63; Indels 41; Gaps 3;


```
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_kin_actsite.
DR Pfam; PF00069; pkinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 566 AA; 64052 MW; 5D57595550902EA9 CRC64;

Query Match 51.7%; Score 941.5; DB 6; Length 566;
Best Local Similarity 49.7%; Pred. No. 1.7e-74;
Matches 177; Conservative 67; Mismatches 65; Indels 47; Gaps 6;

QY 1 MEKYEKLAKTGE GSYGVVFKCRNKTSGQVAVKFKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEKYENLGLVGESYGMVMKCRNKDSGRIVAIAKKFLESDDDKMVKKIAMREIKLLKQLRH 60

QY 61 PNLVNLIEVFRKRKMHLPVEYCDHTLLNELEARNPNGVADGVTKSVLWQTLQALNFCIH 120
Db 61 ENLVNLEVCCKKKRWYLVFEFVDHTILDLELFPNGLDDQVVQKYLFIINGIGFCHSH 120

QY 121 NCIIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDYQYG 179
Db 121 NIIHRDIKPENILVSQGVVKLCDFGFARTLAAPGEVYTDYVATRWYRAPELLVGDVKYG 180

QY 180 SSVDIWAIGCVFAELLTGQPLWPGKSDVDQYLIIRTLIVETGFRHVDQAGLELLTSSDPP 239
Db 181 KAVDVWAIGCLVTEMLMGEPLFPGDSIDQYLIIMRCL----- 218

QY 240 AVASQSAGITGKLI PRHQSIKSNGFHGISIPEPEMETLEEKFSVDVHPVALNFMKGCL 299
Db 219 -----GNLIPRHQELFYKNPVPFAGVRLPEIKESPLERRYPKLSVVVIDLAKKCL 268

QY 300 KMNPDRLTCSQLLESSYF--DSF-----QEAQIK--RKARN-----EGRNRRRQ 341
Db 269 HVDPKRPFCALLHHDFQMDGFAERFSQELQMKVQKDARNISLSKKSQNRKKEK 324

RESULT 4
Q92772 PRELIMINARY; PRT; 493 AA.
AC Q92772;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE P56 KKIAMRE PROTEIN KINASE.
GN KKIAMRE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97152547; PubMed=9000130;
RA Taglienti C.A., Wysk M.A., Davis R.J.;
RT "Molecular cloning of the epidermal growth factor-stimulated protein
kinase p56 KKIAMRE.";
RT Oncogene 13:2563-2574(1996).
RL -!- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC EMBL; U35146; AAC50918.1; -.
DR HSSP; P24941; 1HCL.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_kin_actsite.
DR Pfam; PF00069; pkinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 493 AA; 56018 MW; 5CC20A91CBF89EFE CRC64;
```

```
Query Match 50.3%; Score 915.5; DB 4; Length 493;
Best Local Similarity 48.3%; Pred. No. 2.8e-72;
Matches 172; Conservative 68; Mismatches 69; Indels 47; Gaps 6;

QY 1 MEKYEKLAKTGE GSYGVVFKCRNKTSGQVAVKFKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEKYENLGLVGESYGMVMKCRNKDTGRIVAIAKKFLESDDDKMVKKIAMREIKLLKQLRH 60

QY 61 PNLVNLIEVFRKRKMHLPVEYCDHTLLNELEARNPNGVADGVTKSVLWQTLQALNFCIH 120
Db 61 ENLVNLEVCCKKKRWYLVFEFVDHTILDLELFPNGLDYQVVQKYLFIINGIGFCHSH 120

QY 121 NCIIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDYQYG 179
Db 121 NIIHRDIKPENILVSQGVVKLCDFGFARTLAAPGEVYTDYVATRWYRAPELLVGDVKYG 180

QY 180 SSVDIWAIGCVFAELLTGQPLWPGKSDVDQYLIIRTLIVETGFRHVDQAGLELLTSSDPP 239
Db 181 KAVDVWAIGCLVTEMLMGEPLFPGDSIDQYLIIMMCL----- 218

QY 240 AVASQSAGITGKLI PRHQSIKSNGFHGISIPEPEMETLEEKFSVDVHPVALNFMKGCL 299
Db 219 -----GNLIPRHQELFNKNPVPFAGVRLPEIKERPLERRYPKLSVVVIDLAKKCL 268

QY 300 KMNPDRLTCSQLLESSYF--DSF-----QEAQIK--RKARN-----EGRNRRRQ 341
Db 269 HIDPKRPFCALLHHDFQMDGFAERFSQELQMKVQKDARNVSLSKSQNRKKEK 324

RESULT 5
Q9QYI2 PRELIMINARY; PRT; 564 AA.
AC Q9QYI2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE SER/THR KINASE KKIAMRE-GAMMA.
GN CDKL2 OR KKM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=BRAIN;
RA Sassa T., Gomi H., Sun W., Ikeda T., Thompson R.F., Itohara S.;
RT "The Murine KKIAMRE gene; variants, dual promoters, expression and
chromosomal localization.";
RT Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB029066; BAA88428.1; -.
DR HSSP; P27703; 1ERK.
DR MGD; MGI:1858227; Cdkl2.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_kin_actsite.
DR Pfam; PF00069; pkinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 564 AA; 63640 MW; DACDBF630CCE82D5 CRC64;
```

Query Match 49.7%; Score 904.5; DB 11; Length 564;
Best Local Similarity 48.6%; pred. No. 3.2e-71;
Matches 174; Conservative 66; Mismatches 69; Indels 49; Gaps 7;

QY 1 MEKYEKLAKTGE GSYGVVFKCRNKTSGQVAVKFKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEKYENLGLVGESYGMVMKCRNKDSGRIVAIAKKFLESDDDKMVKKIAMREIKLLKQLRH 60

DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Kinase; Transferase.
SQ SEQUENCE 1106 AA; 118884 MW; 24AC26F5C209923E CRC64;

Query Match 40.9%; Score 743.5; DB 5; Length 1106;
Best Local Similarity 43.7%; Pred. No. 1.3e-56;
Matches 145; Conservative 62; Mismatches 88; Indels 37; Gaps 3;

QY 1 MEKYEKLAKTGECSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEAYETLGIIGEGTYGVVVKARSVRTGKLVAKRFKQTEQDEHVRKTSRSREVRMLQLQH 60

QY 61 PNLVNLIEVFRKRKMHLVFEYCDHTLLNELEARNPNPGVADGVIKSVLWOTLQALNFCCHI 120
Db 61 PNVRLEDVFRREGKLYLVFEFIDHTILQLLESTRGFHRHELRRYTYQLLRGIEFCHNQ 120

QY 121 NCIHEDIKPENILITKQGIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDQYGS 180
Db 121 NIHRDVKPENVLIDESGLLKLCDFGFARQTSAGKQYTDYVATRWYRAPELLVGDVAYGK 180

QY 181 SVDIWAIGCVFAELLTGQPLWPKSDVDQLYLIIRTLETGFRHVDQAGLELLTSSDPPA 240
Db 181 PVDVWALGCMFAELSDGQPLFPGESDLQCLIMQT----- 216

QY 241 VASQSAGITGKLIPRHQSIFKSNGFHGISIPEPEMETLEEKFSVHPVALNFMKGCLK 300
Db 217 -----CGVPQRLVFIEMHNPLYNGISFPHTDILYTLKDRYHRESNDWIEFLSSCLH 268

QY 301 MNPDRLTCSQLLESSYF--DSFQ---EAQIK 327
Db 269 TDPAQRLTCTELMELPYFTRDGRFDRYEAELR 300

RESULT 11
Q9JM02
ID Q9JM02 PRELIMINARY; PRT; 457 AA.
AC Q9JM02;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE SERINE/THREONINE KINASE NKIATRE BETA.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Haq R., Randall S., Midmer M., Iafrate B., Zanke B.W.;
RT "Cloning and functional characterization of NKIATRE, a novel kinase
RT related to both MAPKs and cyclin-dependent kinases.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF112183; AAF34870.1; -.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_kin_actsite.
DR Pfam; PF00069; pkinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 457 AA; 51894 MW; 99D0782794099CDA CRC64;

Query Match 40.5%; Score 737.5; DB 11; Length 457;
Best Local Similarity 42.5%; Pred. No. 1.2e-56;
Matches 145; Conservative 65; Mismatches 92; Indels 39; Gaps 5;

QY 1 MEKYEKLAKTGECSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEAYETLKGKVGESYGVVVKCKHKDTGRIVAIFKIFYE-KPEKSVNKIATREIKFLKQFRH 59

QY 61 PNLVNLIEVFRKRKMHLVFEYCDHTLLNELEARNPNPGVADGVIKSVLWOTLQALNFCCHI 120
Db 60 ENLVNLIEVFRKOKKIHLVFEFIDHTVLDELQHYCHGLESKRRLKYLFOILRAIEYLNHN 119

QY 121 NCIHEDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDQYGS 179
Db 120 NIHRDIKPENILVSQSGITKLCDFGFARTLAAPGDVYTDYVATRWYRAPELLVKDITYG 179

QY 180 SSVDIWAIGCVFAELLTGQPLWPKSDVDQLYLIIRTLETGFRHVDQAGLELLTSSDPP 239
Db 180 KPVDIWAIGCMIEMATGNPYLPSSSDLDLLHKIVLKV----- 217

QY 240 AVASQSAGITGKLIPRHQSIFKSNGFHGISIPEPEMETLEEKFSVHPVALNFMKGCL 299

QY 61 PNLVNLIEVFRKRKMHLVFEYCDHTLLNELEARNPNPGVADGVIKSVLWOTLQALNFCCHI 120
Db 60 ENLVNLIEVFRKOKKIHLVFEFIDHTVLDELQHYCHGLESKRRLKYLFOILRAIEYLNHN 119

QY 121 NCIHEDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDQYGS 179
Db 120 NIHRDIKPENILVSQSGITKLCDFGFARTLAAPGDVYTDYVATRWYRAPELLVKDITYG 179

QY 180 SSVDIWAIGCVFAELLTGQPLWPKSDVDQLYLIIRTLETGFRHVDQAGLELLTSSDPP 239
Db 180 KPVDIWAIGCMIEMATGNPYLPSSSDLDLLHKIVLKV----- 217

QY 240 AVASQSAGITGKLIPRHQSIFKSNGFHGISIPEPEMETLEEKFSVHPVALNFMKGCL 299
Db 218 -----GNLTPHLHNIFSKSPIFAGVVLPOVQHPKNARKKYPKLNGLLADIVHACL 267

QY 300 KMNPDRLTCSQLLESSYF--DSFQ---AQIKRKARNEGR 335
Db 268 QIDPAERISSTDLLHHDYFTRDGFIEKFIPELRAKLLQEA 308

RESULT 12
Q9JM01
ID Q9JM01 PRELIMINARY; PRT; 505 AA.
AC Q9JM01;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE SERINE/THREONINE KINASE NKIATRE ALPHA.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Haq R., Randall S., Midmer M., Iafrate B., Zanke B.W.;
RT "Cloning and functional characterization of NKIATRE, a novel kinase
RT related to both MAPKs and cyclin-dependent kinases.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF112184; AAF34871.1; -.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_kin_actsite.
DR Pfam; PF00069; pkinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 505 AA; 57242 MW; 89747921FF2CE683 CRC64;

Query Match 40.5%; Score 737.5; DB 11; Length 505;
Best Local Similarity 42.5%; Pred. No. 1.4e-56;
Matches 145; Conservative 65; Mismatches 92; Indels 39; Gaps 5;

QY 1 MEKYEKLAKTGECSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEAYETLKGKVGESYGVVVKCKHKDTGRIVAIFKIFYE-KPEKSVNKIATREIKFLKQFRH 59

QY 61 PNLVNLIEVFRKRKMHLVFEYCDHTLLNELEARNPNPGVADGVIKSVLWOTLQALNFCCHI 120
Db 60 ENLVNLIEVFRKOKKIHLVFEFIDHTVLDELQHYCHGLESKRRLKYLFOILRAIEYLNHN 119

QY 121 NCIHEDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDQYGS 179
Db 120 NIHRDIKPENILVSQSGITKLCDFGFARTLAAPGDVYTDYVATRWYRAPELLVKDITYG 179

QY 180 SSVDIWAIGCVFAELLTGQPLWPKSDVDQLYLIIRTLETGFRHVDQAGLELLTSSDPP 239
Db 180 KPVDIWAIGCMIEMATGNPYLPSSSDLDLLHKIVLKV----- 217

QY 240 AVASQSAGITGKLIPRHQSIFKSNGFHGISIPEPEMETLEEKFSVHPVALNFMKGCL 299

```
Db 218 -----GNLTPHLNIFSKSPIFAGVVLPOVQHPKNARKKYPKLNGLLADIVHACL 267
QY 300 KNPDDRLTCSQLLESSYF--DSFOE---AQIKRKARNEGR 335
Db 268 QIDPAERISSDLLLLHHDYFTRDGFIEKFIPELRAKLLOEAK 308

RESULT 13
Q9W6R6
ID Q9W6R6 PRELIMINARY; PRT; 1104 AA.
AC Q9W6R6;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE SERINE-THREONINE KINASE 9.
GN STK9.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99263230; PubMed=10330123;
RA Brunner B., Todt T., Lenzner S., Stout K., Schulz U., Ropers H.-H.,
RA Kalscheuer V.M.;
RT "Genomic structure and comparative analysis of nine Fugu genes:
RT conservation of synteny with human chromosome Xp22.2-p22.1.";
RL Genome Res. 9:437-448(1999).
CC -!- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF146687; AAD28798.1; -.
DR HSSP; Q16539; lwfc.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_kin_actsite.
DR Pfam; PF00069; pkinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 1104 AA; 123751 MW; 9085A078B551d876 CRC64;
```

```
Query Match 34.0%; Score 618; DB 13; Length 1104;
Best Local Similarity 38.0%; Pred. No. 1.4e-45;
Matches 128; Conservative 66; Mismatches 97; Indels 46; Gaps 7;

QY 24 KTSQGVAVKKESEDDPVVKKIALREIRMLKQLKHPNLNLIIEVFRKRKMHLVFEYC 83
Db 19 KETNELVAIKKPKDSEENEVKEETTLRELKMLRTLKQDNIVELKEAFRRGKLYLVEYV 78

QY 84 DHTLNELERNPNGVADGVIKSVLWQTLQALNFCIHNCIHRDIKPENILITKQGIKIC 143
Db 79 ERNMLELEEHPGTGAPPDKVRSYIYQLIKAINWCHKNEIVHRDIKPENLITSSDDILKLC 138

QY 144 DFGFAQILIPG-DA-YTDYVATRWYRAPPELLVGDTOYGSSVDIWAIGCVFAELLTGQPLW 201
Db 139 DFGFARNLSEGTDANYTEYVATRWYRSPELLLG-APYKAVDMWSVGCILGELSDGQPLF 197

QY 202 PGKSDVDQLYLIIRTLVETGFRHVDQAGLELLTSSDPPAVASQSAGITGKLIPRHQSIFK 261
Db 198 PGESEIDQLFTIQKVL-----GPLPAEQMKLFY 225

QY 262 SNGFFHCISIPEDMETLEEKFSDV-HPVALNFMKGCLKMNPDRLTCSQLLESSYFDS 320
Db 226 NNPRFHGIRFPSVTHPQTLERRYQGLSGLMDLKMKNLLLNPTERYLTEQSLNHPAFQP 285

QY 321 FOEAQIKRKA----RNEGRNRR-----QQVLPLKS 347
Db 286 LRQVERERAPPASPNNPPRSSKRRKTHHGENTVPTRS 322
```

```
RESULT 14
Q9UJL6
ID Q9UJL6 PRELIMINARY; PRT; 997 AA.
AC Q9UJL6;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE DJ245GI9.2 (SERINE/THREONINE KINASE 9) (FRAGMENT).
GN STK9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bird C.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
CC -!- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; Z92542; CAB38870.1; -.
DR HSSP; P28523; lA60.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR000508; Peptidase_S26.
DR InterPro; IPR002290; Ser_thr_kin_actsite.
DR Pfam; PF00069; pkinase; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00501; SPASE_I_1; UNKNOWN_1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
FT NON_TER 1
SQ SEQUENCE 997 AA; 111939 MW; B922855AED72608D CRC64;
```

```
Query Match 33.8%; Score 615.5; DB 4; Length 997;
Best Local Similarity 39.0%; Pred. No. 2.1e-45;
Matches 124; Conservative 65; Mismatches 90; Indels 39; Gaps 5;

QY 28 QVAVKKESEDDPVVKKIALREIRMLKQLKHPNLNLIIEVFRKRKMHLVFEYCDHTL 87
Db 4 EIVAIIKKPKDSEENEVKEETTLRELKMLRTLKQDNIVELKEAFRRGKLYLVEYVEKNM 63

QY 88 LNELERNPNGVADGVIKSVLWQTLQALNFCIHNCIHRDIKPENILITKQGIKICDFGF 147
Db 64 LELLEEMPNGVPPPEKVKSYIYQLIKAIHWCHKNDIVHRDIKPENLLISHNDVLKLCDFGF 123

QY 148 AQILIPGD--AYTDYVATRWYRAPPELLVGDTOYGSSVDIWAIGCVFAELLTGQPLWPGKS 205
Db 124 ARNLSEGNNANYTEYVATRWYRSPELLLG-APYKGSVDMWSVGCILGELSDGQPLFPGES 182

QY 206 DVQDLYLIIRTLVETGFRHVDQAGLELLTSSDPPAVASQSAGITGKLIPRHQSIFKSNGF 265
Db 183 EIDQLFTIQKVL-----GPLPEQMKLFYSNPR 210

QY 266 FHGISIPEPEDEMTELEEKFSDV-HPVALNFMKGCLKMNPDRLTCSQLLESSYFDSFOEA 324
Db 211 FHGLRFPVAVNHPOSLERRYLGILNSVLLDLMKNLLKLPADRYLTEQCLN---HPTFQTQ 267

QY 325 QIKRKARNEGRNRRRQQV 342
Db 268 RLDDRSPRSRAKRPYHV 285
```

```
RESULT 15
Q9ZRI1
ID Q9ZRI1 PRELIMINARY; PRT; 294 AA.
AC Q9ZRI1;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE P34CDC2.
GN CDC2TAA.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
```

OC Triticeae; Triticum.
 OX NCBI_TaxID=4565;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Dong C., John P.C.L.;
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; U23409; AAD10483.1; -.
 DR HSSP; P24941; 1JJSU.
 DR Mendel; 37149; Triae; 2321; 37149.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR002290; Ser_thr_kin_actsite.
 DR Pfam; PF00069; pkinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Serine/threonine-protein kinase; Transferase
 SQ SEQUENCE 294 AA; 33996 MW; 77ABC487FA862A72 CRC64;

Query Match	31.1%;	Score 565.5;	DB 10;	Length 294;
Best Local Similarity	35.8%;	Pred. No. 9.1e-42;		
Matches 119;	Conservative 73;	Mismatches 95;	Indels 45;	Gaps 9;
QY 1	MEKYEKLA	TGEGSYGVVFKCRNKTSGOVAVKKFVESEDDPVVKKIALRIRMLKQLKH	60	
Db	1	MEQYKVEKIGEGTYGVVYKARDRTNETIALKKIRLEQEGVGPSTAIRISLLKEMQH	60	
QY 61	PNLVNLI	EFVRKRMHVLFEYCDHTLLNELERNPN-GVADGVIKSVLWQTLQALNFECHI	119	
Db	61	GNIVKLHDVVHSEKRIWLVFEYLDLKKFMDSCPEFAKSPALIKSYLQILRGVAICH	120	
QY 120	HNCIHRDI	KPENILITKQ-GIIKIDCFGAQIL-IPGDAYTDYVATRWYRAPELLVGD	177	
Db	121	HRVLHRLKPNQLLIDRTNALKLADFLARAFGIPVRTFTEHVVTWYRAPELLGARQ	180	
QY 178	YGSSDIWA	IGCVFAELLTGQPLWPGKSDVDQYLIIRTLVETGFRHVDQAGLELLTSSD	237	
Db	181	YSTPVDVMSVGCIFAEMVNVQKPLPPGDSEIDELFKIFRVLG-----TPNE	225	
QY 238	P--PAVASQ	SAGITKGLIPRHQSTFKSNGFFHCISIP--EPEDMETLEEKFSVDVHPVALN	293	
Db	226	QTWPGVSS-----LPDYKSAF-----PRWQAEADLATV---VPNLEPVGLD	262	
QY 294	FMKGCLKM	NPDRLTCSQLLESSYFDSFQEAQ	325	
Db	263	LLSKMLRFEPNKRRTARQALEHEYFKDMEMVQ	294	

Search completed: May 3, 2002, 11:08:12
Job time: 351 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2002, 11:01:26 ; Search time 22.35 Seconds
(without alignments)
349.380 Million cell updates/sec

Title: US-09-671-050-10
Perfect score: 1820
Sequence: 1 MEKYEKLAKTGEGSYGVVFK.....RKARNEGRNRRRQVLPKLS 347

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTus_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pre3. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	557	30.6	298	2	US-08-874-347-25 Sequence 25, Appl
2	557	30.6	298	2	US-08-969-106-2 Sequence 2, Appl
3	557	30.6	298	3	US-09-093-522-25 Sequence 25, Appl
4	554	30.4	297	1	US-08-176-620A-16 Sequence 16, Appl
5	554	30.4	297	2	US-08-461-985-16 Sequence 16, Appl
6	551.5	30.3	294	2	US-08-874-347-26 Sequence 26, Appl
7	551.5	30.3	294	3	US-09-093-522-26 Sequence 26, Appl
8	547.5	30.1	270	2	US-07-857-224B-31 Sequence 31, Appl
9	545	29.9	297	2	US-08-874-347-24 Sequence 24, Appl
10	545	29.9	297	3	US-09-093-522-24 Sequence 24, Appl
11	541	29.7	297	2	US-08-874-347-23 Sequence 23, Appl
12	541	29.7	297	3	US-09-093-522-23 Sequence 23, Appl
13	540	29.7	274	1	US-08-318-947A-20 Sequence 20, Appl
14	540	29.7	274	2	US-08-795-303-20 Sequence 20, Appl
15	538.5	29.6	300	2	US-08-874-347-10 Sequence 10, Appl
16	538.5	29.6	300	3	US-09-093-522-10 Sequence 10, Appl
17	526	28.9	292	1	US-08-154-915-2 Sequence 2, Appl
18	526	28.9	292	2	US-08-464-517-38 Sequence 38, Appl
19	526	28.9	292	2	US-08-246-361A-38 Sequence 38, Appl
20	526	28.9	292	3	US-08-463-772-38 Sequence 38, Appl
21	526	28.9	292	5	PCT-US93-09945-2 Sequence 2, Appl
22	526	28.9	297	4	US-08-932-787B-21 Sequence 21, Appl
23	526	28.9	297	4	US-08-932-012C-21 Sequence 21, Appl
24	526	28.9	297	4	US-08-888-818C-21 Sequence 22, Appl
25	498.5	27.4	297	2	US-08-874-347-22 Sequence 22, Appl
26	498.5	27.4	297	3	US-09-093-522-22 Sequence 22, Appl
27	493.5	27.1	274	2	US-07-857-224B-30 Sequence 30, Appl

28	489.5	26.9	282	1	US-08-318-947A-19 Sequence 19, Appl
29	489.5	26.9	282	2	US-08-795-303-19 Sequence 19, Appl
30	489	26.9	323	2	US-08-874-347-21 Sequence 21, Appl
31	489	26.9	323	3	US-09-093-522-21 Sequence 21, Appl
32	488.5	26.8	317	1	US-08-463-090B-9 Sequence 9, Appl
33	488.5	26.8	317	2	US-08-874-347-18 Sequence 18, Appl
34	488.5	26.8	317	3	US-09-093-522-18 Sequence 18, Appl
35	486	26.7	324	2	US-08-874-347-20 Sequence 20, Appl
36	486	26.7	324	3	US-09-093-522-20 Sequence 20, Appl
37	475	26.1	274	2	US-07-857-224B-29 Sequence 29, Appl
38	470.5	25.9	298	2	US-08-061-636-3 Sequence 3, Appl
39	470.5	25.9	298	2	US-08-874-347-19 Sequence 19, Appl
40	470.5	25.9	298	3	US-09-093-522-19 Sequence 19, Appl
41	470.5	25.9	298	5	PCT-US94-05268-3 Sequence 3, Appl
42	469.5	25.8	326	2	US-08-306-511A-10 Sequence 10, Appl
43	469.5	25.8	326	2	US-08-893-274-10 Sequence 10, Appl
44	467	25.7	316	1	US-08-403-634-4 Sequence 4, Appl
45	467	25.7	316	4	US-08-913-441B-4 Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-874-347-25
; Sequence 25, Application US/08874347
; Patent No. 5863741
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; TITLE OF INVENTION: CARINII
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/874,347
; FILING DATE: 13-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/055001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070
; TELEFAX: 612-288-9696
; TELEX:
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-874-347-25

Query Match 30.6%; Score 557; DB 2; Length 298;
Best Local Similarity 38.4%; Pred. No. 1.2e-50;
Matches 118; Conservative 54; Mismatches 89; Indels 46; Gaps 6;

QY	1	MEKYEKLAKTGE	SGYGVFKCRNKT	SGQVAVKKFV	SEDDPVVKKIALRE	IRMLKQLKH	60				
Db		: :	: : : :	: : : :	:	: : : :					
	1	MENFOKVEKIG	EGTGVVYKARNKLT	GEVVALKKIRLDT	ETEGVPSTAIR	ISLLKELNH	60				
QY	61	PNLVNLIEYFR	RRKKRMLVFEY	CDHTLLNELE	RNP-NGVADGVI	KSVLWQTLQALNF	CH	119			
Db		: :	: : : :	:	: :	: :					
	61	PNIVKLLDVI	HTENKLYLVEF	LHQDLKKFMD	ASALTGIP	LPIKSYLFLQLQGLAF	CHS	120			
QY	120	HNCIHRDIK	PENILITKQI	IKICDFGAQIL-	IPGDAYTDY	VATRWYRAPEL	LVGD	QY	178		
Db		: : : :	: : : :	:	: :	: : : :					
	121	HRVLHRDLK	QPONLLINTE	GAIKLADFG	LARAFGV	PVRTYTHEV	VTLWYRAPEL	LLGS	YY	180	
QY	179	GSSVDIWA	IGCVFAELLT	GTGPLWPKSD	VDVQLYLI	IRTL-----				217	
Db		: : : :	: : : :	: :	: :	: :					
	181	STAVDIWSL	GCIFAEMVTR	ALFPGDSEID	QLFRIFRTL	GLTDPDEVVW	PGVTSMPDY	KPSF		240	
QY	218	---VETG	FRHV---DQ	AGLELLTSS---	DPPAVASQ	SAGITGKL	IPRHQSIF	KSNGF	FFH	267	
Db		: :	: :	: :	: :	: :	: :	: :			
	241	PKWARODF	SKVVPPLDED	GRSLLSQML	HYDPNKR	ISAKAAL-----				AHPFFQ	287
QY	268	GISTPEP		274							
Db		: :									
	288	DVTKPVP		294							

```

RESULT      2
US-08-969-106-2
; Sequence 2, Application US/08969106
; Patent No. 5986055
; GENERAL INFORMATION:
; APPLICANT: Yang, M.
; APPLICANT: Nandabalan, K.
; APPLICANT: Schulz, V.
; TITLE OF INVENTION: CDK2 INTERACTIONS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/969,106
; FILING DATE: 13-NOV-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7934-057
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-969-106-2

```

Query Match 30.6%; Score 557; DB 2; Length 298;
Best Local Similarity 38.4%; Pred. No. 1.2e-50;
Matches 118; Conservative 54; Mismatches 89; Indels 46; Gaps 6;

QY	1	MEKYBKLAKTGETSGYGWVFKCRNKTSQGVAVKKFVESEDDPVVKKIALREIRMLKQLKH	60
Dd	1	MENFOKEKIGETGYGVVYKARNKLTGEVVALKKIRLDTETEGVPSTAIREISLLKELNH	60
QY	61	PNLVNLIEVFRKRKMHLVFEYCDHTLLNELELRNP-NGVADGVTKSVLWQTLOALNFCHI	119
Dd	61	PNIVKLLDVIHTENKLYLVFEFLHQDLKKFMDASALTGIPPLIKSYLFQLLQGGLAFCHS	120
QY	120	HNCIHRDIKPENILITKQGIKICDFGAQIL-IPGDAYTDYVATRWYRAPELLVGDTQY	178
Dd	121	HRVLHRLDKPQNLLINTEGAIKLADFGLARAFGVPRITYTHEVVTWLWYRAPELLGSYY	180
QY	179	GSSVDIWAICGVFAELTGQPLWPCKSDVDQLYLIRTLL-----	217
Dd	181	STAVDIWSLGICIFAEVMYTRRALFGPDSEIDQLFRIFRTLGTDPDEVWPGVTSMPDYKPSF	240
QY	218	---VETGRHV---DQAGLELLTSS--DPPAVASQSAGITCKLIPIRHOSIFKSNGEFH	267
Dd	241	PKWARQDESKVVPPLDEDGRSLLSQMLHYDPNKRISAKAAL-----AHPPFQ	287
QY	268	GISIPEP	274
Dd	288	DVTKPVP	294

```

RESULT      3
US-09-522-25
; Sequence 25, Application US/09093522
; Patent No. 6015700
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; TITLE OF INVENTION: CARINII
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/093,522
; FILING DATE: 08-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/874,347
; FILING DATE: 13-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/055002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070
; TELEFAX: 612-288-9696
; TELEX:
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-522-25

```


MOLECULE TYPE: protein
US-08-461-985-16

Query Match 30.4%; Score 554; DB 2; Length 297;
Best Local Similarity 40.5%; Pred. No. 2.4e-50;
Matches 118; Conservative 58; Mismatches 73; Indels 42; Gaps 7;
QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVVAVKKF-VESEDDPVVYKIALREIRMLKQLK 59
Db 1 MEDYTKIEKIGEGTYGVYKGRHKTGTGQVVAMKKIRLESEEG-VPSTAIRISLLKELR 59
QY 60 HPNLVNLIIEVFRRRKRMHLVFEYCDHTLLNELERNPENG--VADGVIKSVLWQTLQALNFC 117
Db 60 HPNIVSLQDVLMDQSRLLYLIPEFLSMDLKKYLDSPPGQYMDSSLVKSYLEQILQGIVFC 119
QY 118 HIHNCIHRDIKPENILITKOGIIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGD 176
Db 120 HSRVLRDLKPNQLLIDDKGTIKLADFGLARAFGIPRVYTHYVVTWLTWRSPEVLGSA 179
QY 177 QYGSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTL----- 217
Db 180 RYSTPVDIWSIGTIFAELATKKPLFHDGSEIDQLFRIFRALGTPNNEVWPEVESLQDYKN 239
QY 218 -----VETGFRHVDQAGLELLTSS--DPPAVASQSAGITGKLIPRH 256
Db 240 TFPKWKPGSLASHVKNLDENGLDLSKMLIYDP-----AKRISGKMALNH 284

RESULT 6
US-08-874-347-26
; Sequence 26, Application US/08874347
; Patent No. 5863741
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: 13-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/055001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070
; TELEFAX: 612-288-9696
; TELEX:
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 294 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-874-347-26

Query Match 30.3%; Score 551.5; DB 2; Length 294;
Best Local Similarity 34.8%; Pred. No. 4.4e-50;
Matches 116; Conservative 68; Mismatches 102; Indels 47; Gaps 8;
QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVVAVKKFVESEDDPVVYKIALREIRMLKQLKH 60
Db 1 MEQYEKEEKIGEGTYGVVYRARDKVTNETIALKKIRLEQDEGVPSTAIRISLLKEMHH 60
QY 61 PNLVNLIEVFRRRKRMHLVFEYCDHTLLN-----ELERNPNGVADGVIKSVLWQTLQAL 114
Db 61 GNVLRHLDVHSEKRIYLVFEYLDLDLKKFMDSCPEFAKNPT-----LIKSYLYQILRGV 115
QY 115 NFCHIHNCIHRDIKPENILITKQ-GIIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELL 172
Db 116 AYCHSHRVLHRDLKPNQLLIDRRTNALKLADFGLARAFGIPVTFTHVVTWLTWYRAPEIL 175
QY 173 VGDYQYGSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLVETGFRHVDQAGLEL 232
Db 176 LGSRYSTPVDMMWSVGCIFAEMVNQKPLFPGDSEIDELFKIFRVLTGP----- 223
QY 233 LTSSDPPAVASQSAGITGKLIPRHOSIFKSNGFHGISIPEPEDMETLEEKESDVHPVAL 292
Db 224 -NEQSWPGVSS-----LPDYKSAFPK-----WQAQDLATI---VPTLDPAGL 261
QY 293 NPMKGCLKMNPPDDLTCSSQLLESSEYFDSFQEAQ 325
Db 262 DLLSKMLRYEPNKRITARQALEHEYFKDLEMVQ 294

RESULT 7
US-09-093-522-26
; Sequence 26, Application US/09093522
; Patent No. 6015700
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: 08-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/874,347
; FILING DATE: 13-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/055002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070
; TELEFAX: 612-288-9696
; TELEX:
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 294 amino acids

; TELEPHONE: 612-335-5070
; TELEFAX: 612-288-9696
; TELEX:
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 297 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-874-347-24

Query Match 29.9%; Score 545; DB 2; Length 297;
Best Local Similarity 40.2%; Pred. No. 2.2e-49;
Matches 117; Conservative 58; Mismatches 74; Indels 42; Gaps 7;

QY 1 MEKYEKLAKTGEISYGVVFKCRNKTSGQVVAVKKE-VESEDDPVVKKIALREIRMLKQLK 59
||| | : | ||| : ||| : | ||| : ||| : ||| : | ||| : ||| : ||| : | ||| : ||| :
Db 1 MEDYTKIEKIGEGTYGVVYKGRHKTTGQVVAMKKIRLESEEEG-VPSTAIRISLLKELR 59
||| | : | ||| : ||| : | ||| : ||| : ||| : | ||| : ||| : ||| : | ||| : ||| :
QY 60 HPNLVNLIEVFRKRKMHVPEYCDHTLLNELELRPNP--VADGVIKSVLWQTLQALNFC 117
||| | : | ||| : ||| : | ||| : ||| : ||| : | ||| : ||| : ||| : | ||| : ||| :
Db 60 HPNIVSLQDVLMDQSRLLYLFELFSLMDLKKYLDSPPGQYMDSSLVKSLYLQIQIVFC 119
||| | : | ||| : ||| : | ||| : ||| : ||| : | ||| : ||| : ||| : | ||| : ||| :
QY 118 HIHNCIHRDIKPNILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGD 176
| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 120 HSRRLVHRDLKPQNLLIDDKGTIKLADFGLARAFGPIRVYTHEVVTWLYRSPEVLGSA 179
| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 177 QYSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTL----- 217
: | : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 180 RYSTPVDIWSIGTIFAELATKKPLFHGDSEIDQLFRIFRALGTPNNEVWPEVESLQDYKN 239
: | : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 218 -----VETGFRHVDQAGLELITSS---DPPAVASQSAGITGKLIPIRH 256
: | : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 240 TFPKWKPGSLASHVKNLDENGLDLSKMLIYDP-----AKRISGKMALNH 284
: | : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

RESULT 10
US-09-093-522-24
; Sequence 24, Application US/09093522
; Patent No. 6015700
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; TITLE OF INVENTION: CARINII
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/093,522
; FILING DATE: 08-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/874,347
; FILING DATE: 13-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/055002

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070
; TELEFAX: 612-288-9696
; TELEX:
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 297 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-093-522-24

Query Match 29.9%; Score 545; DB 3; Length 297;
Best Local Similarity 40.2%; Pred. No. 2.2e-49;
Matches 117; Conservative 58; Mismatches 74; Indels 42; Gaps 7;

QY 1 MEKYEKLAKTGEISYGVVFKCRNKTSGQVVAVKKE-VESEDDPVVKKIALREIRMLKQLK 59
||| | : | ||| : ||| : | ||| : ||| : ||| : | ||| : ||| : ||| : | ||| : ||| :
Db 1 MEDYTKIEKIGEGTYGVVYKGRHKTTGQVVAMKKIRLESEEEG-VPSTAIRISLLKELR 59
||| | : | ||| : ||| : | ||| : ||| : ||| : | ||| : ||| : ||| : | ||| : ||| :
QY 60 HPNLVNLIEVFRKRKMHVPEYCDHTLLNELELRPNP--VADGVIKSVLWQTLQALNFC 117
||| | : | ||| : ||| : | ||| : ||| : ||| : | ||| : ||| : ||| : | ||| : ||| :
Db 60 HPNIVSLQDVLMDQSRLLYLFELFSLMDLKKYLDSPPGQYMDSSLVKSLYLQIQIVFC 119
||| | : | ||| : ||| : | ||| : ||| : ||| : | ||| : ||| : ||| : | ||| : ||| :
QY 118 HIHNCIHRDIKPNILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGD 176
| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 120 HSRRLVHRDLKPQNLLIDDKGTIKLADFGLARAFGPIRVYTHEVVTWLYRSPEVLGSA 179
| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 177 QYSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTL----- 217
: | : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 180 RYSTPVDIWSIGTIFAELATKKPLFHGDSEIDQLFRIFRALGTPNNEVWPEVESLQDYKN 239
: | : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 218 -----VETGFRHVDQAGLELITSS---DPPAVASQSAGITGKLIPIRH 256
: | : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 240 TFPKWKPGSLASHVKNLDENGLDLSKMLIYDP-----AKRISGKMALNH 284
: | : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

RESULT 11
US-08-874-347-23
; Sequence 23, Application US/08874347
; Patent No. 5863741
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; TITLE OF INVENTION: CARINII
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/874,347
; FILING DATE: 13-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/055001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070

Fri May 3 11:22:24 2002

Query Match		29.6%;	Score 538.5;	DB 2;	Length 300;		
Best Local Similarity		34.6%;	Pred. No. 1.1e-48;				
Matches 113;		Conservative 76;	Mismatches 95;	Indels 43;	Gaps 9;		
QY	1	MEKYEKLAKTGE	SGYGVVFKCRNKTSGQVAVKKEVSEDDPVVKKIALREIRMLKQLKH	60			
Db	1	MEQYQRLEKIGEGTYGVVYKAKDLE	SGTIVALKKIRLEAEDEGVPSTAIRISLLKEMHN	60			
QY	61	PNLVNLI	EVFRRKRKMHVFEYCDHTLLNELERNPNGVADG--VIKSVLWQTLQALNFCH	118			
Db	61	DNVVRLLNIIHQESRLYL	VFEFLDLDLKKYMN	SIPKDMMLGAEMIKKFMSQLVSGVKYCH	120		
QY	119	IHNCIHRDIK	PENILITKOGIIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGD	TQ	177		
Db	121	SHRILHRDLK	PQNLLIDREGNKLADFG	LARAFGVPLRGYTHEVVTLWYRAPEVLLGGRQ	180		
QY	178	YGSSVDI	WAIGCVFAELLTGQPLWPGKSDVDQ	LYLIIRTLVETGFRHVDQAGLELLTSSD	237		
Db	181	YATALDIWSIGC	IPAEEMATKKPLFPGDSEIDEIFRIFR-----ILGTPD	224			
QY	238	PPAVASQ	SAGITGKLI	PRHQSI	PKSNGFFHGISIPE--PEDM-ETLEEKFS	DVHPVALNF	294
Db	225	----	ENSWPGITS--YPDFKATF-----PKNSPKNLGELITELDSD	----	GIDL	263	
QY	295	MKGCLK	MNPDDRLTCSQLLESSYFDSF	321			
Db	264	LQKCLRY	YPAERISAKKALDHPYFDDF	290			

Search completed: May 3, 2002, 11:02:44
Job time: 78 sec

Fri May 3 11:02:07 2002

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 16:29:21 ; Search time 2919.35 Seconds
(without alignments)
5340.173 Million cell updates/sec

Title: US-09-671-050-11
Perfect score: 945
Sequence: 1 atggaaaagtatgaaaaatt.....agggtacttcgcgtcaaaagt 945

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rod:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rod:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	945	100.0	945	6	AX107722	AX107722 Sequence
2	929	98.3	972	6	AX107716	AX107716 Sequence
3	863	91.3	1083	6	AX166534	AX166534 Sequence
4	839	88.8	1041	6	AX107720	AX107720 Sequence
5	823	87.1	1068	6	AX107714	AX107714 Sequence
6	643	68.0	911	6	AX056404	AX056404 Sequence
7	559.2	59.2	2615	6	AX056405	AX056405 Sequence
8	507	53.7	1819	6	AX107724	AX107724 Sequence
9	454.4	48.1	561	6	AX107712	AX107712 Sequence
10	454.4	48.1	594	6	AX107718	AX107718 Sequence
11	413.8	43.8	1363	9	HSSTHPKB	X66358 H.sapiens m
12	287	30.4	3153	4	AB029045	AB029045 Oryctolag
13	283.8	30.0	1993	9	HSU35146	U35146 Human p56 K
14	272.6	28.8	1518	10	AB029067	AB029067 Mus muscu
15	272.6	28.8	3988	10	AB029066	AB029066 Mus muscu
16	272.6	28.8	4022	10	AB029065	AB029065 Mus muscu
17	227.4	24.1	1513	9	AF130372	AF130372 Homo sapi
18	227.4	24.1	1773	6	AX166533	AX166533 Sequence
19	218.4	23.1	101340	2	AC018104	AC018104 Drosophil
20	218.4	23.1	190668	3	AC007419	AC007419 Drosophil
21	218.4	23.1	259973	3	AE003611	AE003611 Drosophil
22	211.4	22.4	1732	10	AF112183	AF112183 Rattus no
23	211.4	22.4	1738	10	AF112184	AF112184 Rattus no
24	208.4	22.1	1974	3	AF326965	AF326965 Trypanoso
25	200.4	21.2	80362	9	AC079615	AC079615 Homo sapi
26	200.4	21.2	155028	33	AC037489	Ac037489 Homo sapi
27	200.4	21.2	164281	2	AC092672	AC092672 Homo sapi
28	194	20.5	3399	9	HSTXP3	Y15057 Homo sapien
29	180.2	19.1	3321	3	LME293284	AJ293284 Leishmani
30	175.8	18.6	1866	6	AX056407	AX056407 Sequence
31	171.8	18.2	903	8	AF026545	AF026545 Pneumocys
32	168	17.8	960	3	TCU69958	U69958 Trypanosoma
33	163.6	17.3	921	3	DDICRP	L00652 Dictyosteli
34	163.4	17.3	1476	6	AR087351	AR087351 Sequence
35	163.4	17.3	1476	9	HSCDK2MR	X61622 H.sapiens C
36	163.4	17.3	1635	6	A84497	A84497 Sequence 11
37	163.4	17.3	1635	6	A84499	A84499 Sequence 11
38	163	17.2	1297	9	HSCDK2	X62071 H.sapiens C
39	161.8	17.1	2213	9	HUMCDC2A	M68520 Human cdc2-
40	161.8	17.1	2264	9	BC003065	BC003065 Homo sapi
41	157.2	16.6	885	3	AF076534	AF076534 Cryptospo
42	157	16.6	2213	10	CGR223949	AJ223949 Cricetulu
43	155.6	16.5	1050	9	HSCDC2	X05360 Human CDC2
44	155.6	16.5	1050	9	HSCDC2R	Y00272 Human cell
45	153.8	16.3	1708	10	MMU63337	U63337 Mus musculu

ALIGNMENTS

RESULT 1	AX107722	AX107722	945 bp	DNA	PAT	30-APR-2001
LOCUS	Sequence 11	from Patent	WO0123579.			
DEFINITION	AX107722					
ACCESSION	AX107722					
VERSION	AX107722.1	GI:13923202				
KEYWORDS	human.					
SOURCE	Homo sapiens					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1 (bases 1 to 945)					
AUTHORS	Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.					
TITLE	Human kinase proteins and polynucleotides encoding the same					
JOURNAL	Patent: WO 0123579-A 11 05-APR-2001;					
FEATURES	Lexicon Genetics Incorporated (US)					
SOURCE	Location/Qualifiers					
	1..945					
	/organism="Homo sapiens"					

BASE COUNT		/db_xref="taxon:9606"			
ORIGIN	323 a	165 c	202 g	255 t	
Query Match 100.0%; Score 945; DB 6; Length 945;					
Best Local Similarity 100.0%; Pred. No. 2.2e-202;					
Matches 945; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	atggaaaagtatgaaaaattagctaagactggagaagggtcttattatgggggttgattcaaa	60		
Db	1	ATGGAAAAGTATGAAAAATTAGCTAAGACTGGAGAAGGGTCTTATGGGGTTGTATTCAA	60		
QY	61	tgcagaaaacaaacacctctggacaagtagtctgtttaaaaaattttgtggaatctgaagat	120		
Db	61	TGCAGAAAACAAACCTCTGGACAAGTAGTACTGTTAAAAAATTTTGTGGAATCTGAAGAT	120		
QY	121	gatcctgtgtttaagaaaaatagcactaagagaaatacgtlatgttggaagcaattaaaaacat	180		
Db	121	GATCCTGTGTTAAGAAAAATAGCAATAAGCAAAATACGTATGTTGAAGCAATTAAAACAT	180		
QY	181	ccaaatcttgtgaacctcategaggtgttcaggagaaaaaggaaaaatgcatttagttttt	240		
Db	181	CCAAATCTTGTGAACCTCATCGAGGTGTTCAAGGAGAAAAAGGAAAAATGCATTAGTTTTT	240		
QY	241	gaatacttgtatcatcacacttttaaatgagctggagaaagaaacccaaatggagttgctgat	300		
Db	241	GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT	300		
QY	301	ggagtgatcaaaagcgtattatggcaaacacttcaagctcttaatttctgtcatatacat	360		
Db	301	GGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTCATATACAT	360		
QY	361	aactgtattcacagagatatataaacctgaaaaatatcttaataactaagcaaggaaataatc	420		
Db	361	AACGTATTACAGAGATATAAAACCTTGAAAAATATTCTAATAACTAAGCAAGGAATAATC	420		
QY	421	aagatttgtgacctcgggtttgcacaaattctgattccaggagatgcctacacccgattat	480		
Db	421	AAGATTGTGACTTCGGGTTTGCACAAATCTGATTCAGGAGATGCCATACCCGATTAT	480		
QY	481	gtagctacgagatggtacgagctcctgaacttcttgtggagatactcagttggttct	540		
Db	481	GTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTTGTGGAGATACTCAGTATGGTTCT	540		
QY	541	tcagtcgatataatggcgtattgtgtgttttgcagagctcctgacagggccagccactg	600		
Db	541	TCAGTCGATATATGGGCTATTGGTGTGTTTGTTCAGAGCTCCTGACAGGCCAGCCACTG	600		
QY	601	tggcctggaaaatcagatgtggacaaactttatctgataatcagaaacactaggaaaaatta	660		
Db	601	TGGCCTGGAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGAAAAATTA	660		
QY	661	atcccaagacatcaatcaatctttaaaagtaacgggttttccatggcatcagttacacct	720		
Db	661	ATCCCAAGACATCAATCAATCTTTAAAAGTAACGGGTTTTTCCATGGCATCAGTATACCT	720		
QY	721	gagccagaagacatggaactcttgaggaagaagtctcagatgtttcatcctgtggctctg	780		
Db	721	GAGCCAGAAGACATGGAAACTCTTGAGGAAAAGTTCTCAGATGTTTCATCCTGTGGCTCTG	780		
QY	781	aacttcatgaagggtgtctgaagatgaatccagatgacagattaaacctgttcccaactc	840		
Db	781	AACTTCATGAAGGGGTGCTGAAGATGAATCCAGATGACAGATTAAACCTGTTCCCAACTC	840		
QY	841	ctggagagctcctacttttgattcttttcaagaggccaaaaattaaaaagaaaaagcacgtaat	900		
Db	841	CTGGAGAGCTCCTACTTTTGATTCCTTTTCAAGAGGCCCAAATTAAGAAAAAGCACGTAAT	900		
QY	901	gaaggagaacacagaagacgccaacaggtaacttccgctcaaaagt	945		
Db	901	GAAGGAAGAAACAGAGACGCCAACAGGTACTTCCGCTCAAAAGT	945		

RESULT 2									
AX107716	LOCUS	AX107716	972 bp	DNA	PAT				
DEFINITION		Sequence 5 from Patent WO0123579.		30-APR-2001					
ACCESSION		AX107716							
VERSION		AX107716.1		GI:13923199					
KEYWORDS		human.							
SOURCE		Homo sapiens							
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.							
REFERENCE		1 (bases 1 to 972)							
AUTHORS		Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.							
TITLE		Human kinase proteins and polynucleotides encoding the same							
JOURNAL		Patent: WO 0123579-A 5 05-APR-2001;							
FEATURES		Lexicon Genetics Incorporated (US)							
source		Location/Qualifiers							
		1..972							
		/organism="Homo sapiens"							
		/db_xref="taxon:9606"							
BASE COUNT		323 a	172 c	214 g	263 t				
ORIGIN									
Query Match 98.3%; Score 929; DB 6; Length 972;									
Best Local Similarity 100.0%; Pred. No. 8.7e-199;									
Matches 929; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	atggaaaagtatgaaaaattagctaaagactggagaaagggtcttattgggggttgattc	60						
Db	1	ATGGAAGTATGAAAAATTAGCTAAGACTGGAGAAGGGTCTTATGGGGTTGTATTCAA	60						
QY	61	tgcagaaaacaaacacctctgcagcaagtagtctgttaaaaaatttgggaatctgaagat	120						
Db	61	TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT	120						
QY	121	gatcctgttgttaagaaaaatagcactaagagaaatacgtatgttgaagcaattaaaaacat	180						
Db	121	GATCCTGTGTTAAGAAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACACAT	180						
QY	181	ccaaatcttgtgaacctcctcaggtgtgttcaggagaaaaagggaaaaatgcatttagtttt	240						
Db	181	CCAAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAAAATGCATTAGTTTTT	240						
QY	241	gaatacttgtatcatcacacttttaaatgagctggagaaagaaacccaaatggagttgctgat	300						
Db	241	GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT	300						
QY	301	ggagtgatcaaaagcgtattatggcaaacacttcaagctcttaatttctgtcatatacat	360						
Db	301	GGAGTGATCAAAAGCGTATTATGCAAAACACTTCAAGCTCTTAATTTCTGTCATATACAT	360						
QY	361	aactgtattcacagagatatataaacctgaaaaatatcttaataactaagcaaggaaataatc	420						
Db	361	AACGTATTACAGAGATATAAAACCTGAAAATATTCTAATAACTAAGCAAGGAATAATC	420						
QY	421	aagatttgtgacctcgggtttgcacaaattctgattccaggagatgcctacacccgattat	480						
Db	421	AAGATTGTGACTTCGGGTTTGCACAAATCTGATTCAGGAGATGCCATACCCGATTAT	480						
QY	481	gtagctacgagatggtacccagctcctgaacttcttgtggagatactcagtatggttct	540						
Db	481	GTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTTGTGGAGATACTCAGTATGGTTCT	540						
QY	541	tcagtcgatataatggcgtattgtgtgtgttttgcagagctcctgacagggccagccactg	600						
Db	541	TCAGTCGATATATGGGCTATTGGTGTGTTTTCAGAGCTCCTGACAGGCCAGCCACTG	600						
QY	601	tggcctggaaaatcagatgtggacaaactttatctgataatcagaaacactaggaaaaatta	660						
Db	601	TGGCCTGGAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGAAAAATTA	660						

Db	301	GGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTGCATATACAT	360
QY	361	aactgtattcacagagatatataaacctgaaataatttctaataactaagcaaggaataatc	420
Db	361	AACTGTATTACAGAGATATAAAACCTGAAATATTCTATAACTAAGCAAGGAATAATC	420
QY	421	aagattttgacttcgggtttgcacaaaattctga-----	454
Db	421	AAGATTTGTGACTTCGGGTTTGCACAAATCTCGAGTAATTCACATGTTGGGAGAAATTGAC	480
QY	455	-----ttccaggagatgcctacacagattatgtagct	486
Db	481	TTACCTGATCTGATTGATGCATTTGCAGTTCCAGGAGATGCCACACGATTATGTAGCT	540
QY	487	acgagatggtaccgagctcctgaactcttctgtgggagatactcagtatggttcttcagtc	546
Db	541	ACGAGATGGTACCGAGCTCCTGAACCTCTTGTGGGAGATACTCAGTATGGTTCTTTCAGTC	600
QY	547	gatatatgggctattggtgtgtttttgcagagctcctgacaggccagccactgtggcct	606
Db	601	GATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAGCCACTGTGGCCT	660
QY	607	gaaaaatcagatgtggaccaactttatctgataatacagaacactaggaataataatccca	666
Db	661	GGAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGAAAATTAATCCCA	720
QY	667	agacatcaatcaatcttttaaaagtaacgggtttttccatggcatcagatacactgagcca	726
Db	721	AGACATCAATCAATCTTTAAAGTAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCA	780
QY	727	gaagacatggaactcttgaggaaaagttctcagatgttcatcctgtggtctgaacttc	786
Db	781	GAAGACATGGAACACTCTTGAGGAAAAGTTCTCAGATGTTTCATCCTGTGGCTCTGAAC TTC	840
QY	787	atgaaggggtgtctgaagatgaatccagatgacagatgaattacactgttcccaactcctggag	846
Db	841	ATGAAGGGGTGTCGAAGATGAATCCAGATGACAGATTAACCTGTTCCTCAACTCTCTGGAG	900
QY	847	agctcctactttgattcttttcaagagggcccaaatataaagaaaaagcacgtaataagagga	906
Db	901	AGCTCCTACTTTTGATTCTTTTCAAGAGGGCCCAAAATTAAGAGAAACGACGTAATGAAGGA	960
QY	907	agaaacagaagacgccaacag 927	
Db	961	AGAAACAGAAGACGCCAACAG 981	
RESULT	4		
AX107720		AX107720 1041 bp DNA PAT 30-APR-2001	
LOCUS		Sequence 9 from Patent WO0123579.	
DEFINITION		AX107720	
ACCESSION		AX107720	
VERSION		AX107720.1 GI:13923201	
KEYWORDS		human.	
SOURCE		Homo sapiens	
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE		1 (bases 1 to 1041)	
AUTHORS		Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.	
TITLE		Human kinase proteins and polynucleotides encoding the same	
JOURNAL		Patent: WO 0123579-A 9 05-APR-2001;	
FEATURES		Lexicon Genetics Incorporated (US)	
source		Location/Qualifiers	
		1..1041	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
BASE COUNT		344 a 190 c 228 g 279 t	
ORIGIN			
Query Match		88.8%; Score 839; DB 6; Length 1041;	

QY	661	atcccaagacatcaatcaatctttaaaagtaacgggtttttccatggccatcagtatatacct	720
Db	661	ATCCCAAGACATCAATCAATCTTTAAAGTAACGGGTTTTTCCATGGCATCAGTATACCT	720
QY	721	gagccagaagacatgaaaactcttgaggaaaagttctcagatgttcatcctgtggctctg	780
Db	721	GAGCCAGAAGACATGGAAACTCTTGAGGAAAAGTTCTCAGATGTTTCATCCTGTGGCTCTG	780
QY	781	aacttcataaggggtgtctgaagatgaatccagatgacagattaaacctgttcccaactc	840
Db	781	AAC TTCATGAAGGGGTCTCTGAAGATGAATCCAGATGACAGATTAACTGTTCCTCAACTC	840
QY	841	ctggagagctcctactcttgattcttttcaagaggcccccaattaaaagaaaagcacqtaat	900
Db	841	CTGGAGAGCTCCTACTTTGATTCTTTTCAAGAGGCCCAAAATTAAGAGAAAAGCACGTAAT	900
QY	901	gaaggaagaacagagaagacgcccaacaggt 929	
Db	901	GAAGGAAGAACAAGAGACGCCCAACAGGT 929	
RESULT	3		
AX166534		AX166534 1083 bp DNA PAT 22-JUN-2001	
LOCUS		Sequence 25 from Patent WO0138503.	
DEFINITION		AX166534	
ACCESSION		AX166534	
VERSION		AX166534.1 GI:14546879	
KEYWORDS		human.	
SOURCE		Homo sapiens	
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE		1 (bases 1 to 1083)	
AUTHORS		Plowman,G.D., Whyte,D., Manning,G.S., Sudarsanam,S.S., Martinez,R., Flanagan,P. and Clary,D.S.	
TITLE		Novel human protein kinases and protein kinase-like enzymes	
JOURNAL		Patent: WO 0138503-A 25 31-MAY-2001;	
FEATURES		Sugen, Inc. (US)	
source		Location/Qualifiers	
		1..1083	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
BASE COUNT		366 a 199 c 225 g 293 t	
ORIGIN			
Query Match		91.3%; Score 863; DB 6; Length 1083;	
Best Local Similarity		94.5%; Pred. No. 5.9e-184;	
Matches		927; Conservative 0; Mismatches 0; Indels 54; Gaps 1;	
QY	1	atggaaaagtatgaaaattagctaagactggagaaagggtcttatgggggtgtatttcaaa	60
Db	1	ATGGAAAAGTATGAAAATTAGCTAAGACTGGAGAGGGCTCTTATGGGGTTGTATTCAA	60
QY	61	tcgagaaaaaacctctggacaagtagtagctgttaaaaaatttgggaatctgaagat	120
Db	61	TGCAGAAAACAACCTCTGGACAAGTAGTAGCTGTTAAAAAATTGTGGAATCTGAAGAT	120
QY	121	gatcctgttgttaagaaaatagcactaagaagaaatacgtatgttgaagcaattaaaacat	180
Db	121	GATCCTGTTGTTAAGAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAAACAT	180
QY	181	ccaaaatcttgtgaacctcatcgaggtgttcaggagaaaaagaaaatgcatttagttttt	240
Db	181	CCAAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAAATGCAATTAGTTT	240
QY	241	gaatactgtgatcatacaacttttaaatgagctggagaaagaaaccccaaatggagttgctgat	300
Db	241	GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGAT	300
QY	301	ggagtgatcaaaagcgtattatggcaaacacacttcaagctcttaatttctgtcatatacat	360
Db	301	GGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTGCATATACAT	360

Best Local Similarity 90.8%; Pred. No. 1.4e-178;
Matches 945; Conservative 0; Mismatches 0; Indels 96; Gaps 1;

QY 1 atggaagagtataaaatttagtaagactggagaagggtcttatggggtgtattcaaa 60
Db |
1 ATGGAAGAGTATGAAAAATTAGCTAAGACTGGAGAAGGGTCTTATGGGTGTATTCAA 60

QY 61 tgcagaaacaaacacctctggaacaagtagtagctgtttaaaaaatttggaaatcagaat 120
Db |
61 TGCAGAAACRAAACCTCTGGACAAGTAGAGTGTGTTAAAAAATTGTGGAATCTGAAGAT 120

QY 121 gatcctgtgttaagaaaaatagcaactaagagaatactgtatgttgaagcaattaaaaacat 180
Db |
121 GATCCTGTTGTTAAGAAAAATAGCTAAGAGAAATACGTATGTTGAAGCAATTAAACAT 180

QY 181 ccaaatcttgtgaacctcaatcgaggtgttcaggagaaaaaggaaatgcatttttt 240
Db |
181 CCAAAATCTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAATGSCATTTAGTTT 240

QY 241 gaatactgtgatacatcacacttttaaatgagctggaagaaaccccaatggagtgtgat 300
Db |
241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTGTCTGAT 300

QY 301 ggagtgcataaaagcgtattatggaacacattcaagctcttaattctgtcatacat 360
Db |
301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAAATTTCTGTTCATATACAT 360

QY 361 aactgtattcacagagatataaaacacctgaaaaattcttaataactaagcaaggaataatc 420
Db |
361 AACTGTATTACAGAGATATAAAACCTGAAAAATATTCTAATACTAAGCAAGGAATAATC 420

QY 421 aagatttgtgacttcgggtttgcacaaaattctgattccaggagatgcctacaccgattat 480
Db |
421 AAGATTGTGACTTCGGGTTTGCACAAAATCTGATTCAGGAGATGCCTACACCGATTAT 480

QY 481 gtactacgagatggtaccgagctccctgaacttcttgtgggagatactcagtatggtct 540
Db |
481 GTAGCTACGAGATGGTACCAGCTCCTGAACTTCTGTGGAGATACTCAGTATGGTTCT 540

QY 541 tcagtcgatatatgggctattgggtgtgtttttgacagagctcctgacagccagccactg 600
Db |
541 TCAGTCGATATATGGGCTATTGGTTGTGTTTTTGACAGAGCTCCTGACAGCCAGCCACTG 600

QY 601 tggcctggaaaatcagatgtggaccacactttatctgataatcagaaacact----- 650
Db |
601 TGSCCTGGAARATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGTAGAGACG 660

QY 651 ----- 650

Db 661 GGTTTGGCCATGTTGACCAGGCTGGTCTCGAACTCTTGACGTCAAGTGATCCACCTGCC 720

QY 651 -----aggaataataatcccaagacataatcaatcttt 684
Db |
721 GTAGCCTCTCAAAGTGTGGAAATTACAGGAAAAATTAATCCCAAGACATCAATCAATCTTT 780

QY 685 aaaagtaacgggtttttccatggcatcagatcacctgagccagaagacatggaaactctt 744
Db |
781 AAAAGTAACGGGTTTTTCCATGGCATCAGTATACCTTGAGCTCTGAACTTCATGAAGGGGTGCTGAAG 840

QY 745 gaggaagaagtctcagatgttcactcctgtggtctgtgaacttcattcatgaaggggtgtgaag 804
Db |
841 GAGGAAAAAGTTCTCAGATGTTTCATCCTGTGGCTCTGAACTTCATGAAGGGGTGCTGAAG 900

QY 805 atgaatccagatgacagattaacctgtttcccaactcctggagagctcctactttgattct 864
Db |
901 ATGAATCCAGATGACAGATTAACTGTTCCCAACTCTCTGGAGAGCTCCTACTTTGATTCT 960

QY 865 ttcaagaggcccaattaaagaaaaagcacgtaataaggaagaaacagagacgcgcaa 924
Db |
961 TTTCAGAGGGCCCAAATTAAGAGAAAAAGCACGTAATGAAGGAAGAAACAGAGACGCCAA 1020

QY 925 caggtactccgcctcaaaagt 945
|

Db 1021 CAGGTACTTCGCGCTCAAAAGT 1041

RESULT 5
AX107714

LOCUS AX107714 1068 bp DNA PAT 30-APR-2001

DEFINITION Sequence 3 from Patent WO0123579.

ACCESSION AX107714

VERSION AX107714.1 GI:13923198

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1068)

AUTHORS Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.

TITLE Human kinase proteins and polynucleotides encoding the same

JOURNAL Patent: WO 0123579-A 3 05-APR-2001;

FEATURES Location/Qualifiers

source 1..1068

organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT 344 a 197 c 240 g 287 t

ORIGIN

Query Match 87.1%; Score 823; DB 6; Length 1068;
Best Local Similarity 90.6%; Pred. No. 5.7e-175;
Matches 929; Conservative 0; Mismatches 0; Indels 96; Gaps 1;

QY 1 atggaagagtataaaatttagctaagactggagaagggtcttatggggtgtattcaaa 60
Db |
1 ATGGAAGAGTATGAAAAATTAGCTAAGACTGGAGAAGGGTCTTATGGGTGTATTCAA 60

QY 61 tgcagaaacaaacacctctgacaaagtagtagctgttaaaaaatttgtggaatcctgaagat 120
Db |
61 TGCAGAAACRAAACCTCTGGACAAGTAGAGTGTGTTAAAAAATTGTGGAATCTGAAGAT 120

QY 121 gatcctgtgttaagaaaaatagcactaagagaaatacgtatgttgaagcaattaaaaacat 180
Db |
121 GATCCTGTTGTTAAGAAAAATAGCACTAAGAGAAATACGTAATGTAAGCAATTAACACAT 180

QY 181 ccaaatcttgtgaacctcaatcgaggtgttcaggagaaaaaggaaatgcatttttt 240
Db |
181 CCAAAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAATGCAATTTAGTTT 240

QY 241 gaatactgtgatacatcacacttttaaatgagctggaagaaaccccaatggagtgtgat 300
Db |
241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTGTCTGAT 300

QY 301 ggagtgcataaaagcgtattatggcaaacacttcaagctcttaattctgtcatacat 360
Db |
301 GGAGTGATCAAAAGCGTATTATGCAAAACACTTCAAGCTCTTAAATTTCTTCATATACAT 360

QY 361 aactgtattcacagagatataaaacacctgaaaaattcttaataactaagcaaggaataatc 420
Db |
361 AACTGTATTACAGAGATATAAAACCTGAAAAATATTCTAATAACTAAGCAAGGAATAATC 420

QY 421 aagatttgtgacttcgggtttgcacaaaattctgattccaggagatgcctacaccgattat 480
Db |
421 AAGATTGTGACTTCGGGTTTGACAAATTTCTGATTTCCAGGAGATGCCTACACCGATTAT 480

QY 481 gtactacgagatggttaccgagctcctgaacttcttgtgggagatactcagtatggtct 540
Db |
481 GTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTTGTGGGAGATACTCAGTATGGTTCT 540

QY 541 tcagtcgatatatgggctattgggtgtgtttttgcagagctcctgcagagccagccactg 600
Db |
541 TCAGTCGATATATGGGCTATTGGITGTGTTTTTGCAGAGCTCCTGCACAGGCCAGCCACTG 600

QY 601 tggcctggaaaatcagatgtggaccacactttatctgataatcagaaacact----- 650

Db 601 TGGCCTGGAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGTAGAGACG 660
QY 651 ----- 650
Db 661 GGGTTTCGCCATGTTGACCAGGCTGGFCTCGAACTCTTGACGTCAAAGTCACACCTGCC 720
QY 651 -----agggaaaattaatcccaagacatcaatcaatcttt 684
Db 721 GTAGCCTCTCAAAGTGTCTGGAATTACAGGAAAATTAAATCCCAAGACATCAATCAATCTTT 780
QY 685 aaaagtaacgggtttttccatggcatcagtataacctgagccagaagacatggaaaaacttt 744
Db 781 AAAAGTAACGGGFTTTTCCATGGCATCAGTATACCTGTGCTCGAACTTCATGAAGGGGTGCTGAAG 840
QY 745 gaggaagaagtctcagatgttcacctcctgtgtcctgaaccttcataagagggtgtctgaag 804
Db 841 GAGGAAAAGTTCTCAGATGTTTCATCCTGTGGCTCTGAACTTCATGAAGGGGTGCTGAAG 900
QY 805 atgaatccagatgacagattaacctgttcccaactcctggagagctcctactttgattct 864
Db 901 ATGAATCCAGATGACAGATTTAACTGTGTCCCAACTCCTGGAGAGCTCCTACTTTGATTCT 960
QY 865 ttccaagaggcccaaatataaagaaaaagcacgtaatgaaggaaagaaacagagacgcca 924
Db 961 TTTCAAGAGGCCCAAAATTAAGAAAAGCAGCGTAATGAAGGAAGAAACAGAGACGCCAA 1020
QY 925 caggt 929
Db 1021 CAGGT 1025

RESULT 6
AX056404 AX056404 911 bp DNA PAT 13-JAN-2001
LOCUS Sequence 48 from Patent WO0073469.
DEFINITION AX056404
ACCESSION AX056404
VERSION AX056404.1 GI:12229111
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 911)
AUTHORS Plowman,G.D., Martinez,R., Whyte,D. and Sudersanam,S.
TITLE Protein kinases
JOURNAL Patent: WO 0073469-A 48 07-DEC-2000;
Sugen, Inc. (US)
FEATURES Location/Qualifiers
source 1..911
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 293 a 182 c 178 g 258 t
ORIGIN
Query Match 68.0%; Score 643; DB 6; Length 911;
Best Local Similarity 100.0%; Pred. NO. 1.6e-134;
Matches 643; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 285 aaatggaggttctgtgatggatgatcaaaaagcgtattatggcaaacacttcaagctcttaa 344
Db 1 AAATGGAGTTGCTGATGGAGTGATCAAAGCGTATATATGGCAAAACACTTCAAGCTCTTAA 60
QY 345 ttctgtcatatacataactgtattcacagagatataaaacctgaaaaattcttaataac 404
Db 61 TTTCTGTCATATACATAACTGTATTTCACAGAGATATAAACCTGAAAAATATTCTAATAAC 120
QY 405 taagcaaggaataatacaagatttgtgacttcgggttgcacaaattctgattccaggaga 464
Db 121 TAAGCAAGGAATAATCAAGATTGTGACTTCGGGTTCACAAAATTCGTGATTCAGGAGA 180
QY 465 tgcctacaccgattatgtagctacgagatgggtaccgagctcctgaacttcttctgtgggaga 524

Db 181 TGCTACACCGATTATGTAGCTACGAGATGGTACGAGCTCCTGAACCTCTTGTGGAGA 240
QY 525 tactcagtatggttcttcagtcgatataatgggctattgggtgtgttttgcagagctcct 584
Db 241 TACTCAGTATGGTCTTTCAGTCGATATATGGCTATTGGTTGTGTGTTTTCGAGAGCTCCT 300
QY 585 gacagggccagccactgtggcctggaaaatcagatgtggaccaactttatctgataatacag 644
Db 301 GACAGGCCAGCCACTGTGGCTTGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAG 360
QY 645 aacctaggaaaaattaatcccaagacatcaatcaatcttttaaaagtaacgggtttttcca 704
Db 361 AACACTAGGAAAATTAATCCCAAGACATCAATCAATCTTTAAAAAGTAACGGGTTTTTCCA 420
QY 705 tggcatcagtalacctgagccagaagacatggaaactccttgagaaaaagttctcagatgt 764
Db 421 TGGCATCAGTATACCTGAGCCAGAAGACATGGAAACTCTTTGAGGAAAAGTTCTCAGATGT 480
QY 765 tcatcctgtggctctgaacttcattgaagggtgtctgaagatgaatccagatgacagatt 824
Db 481 TCATCCTGTGGCTCTGAACCTTCATGAAGGGGTGTCTGAAGATGAATCCAGATGACAGATT 540
QY 825 aacctgtcccaactcctggagagctcctactttgattcttttcaagaggcccaaatata 884
Db 541 AACCTGTTCCTCAACTCCTGGACAGCTCCTACTTTGATTCTTTTCAAGAGGGCCCCAAATTAA 600
QY 885 aagaaaagcacgtaatgaaggagaagaacagaagacgccaacag 927
Db 601 AAGAAAAGCACGTAATGAAGGAAGAAACACAGAGACGCCCAACAG 643

RESULT 7
AX056405 AX056405 2615 bp DNA PAT 13-JAN-2001
LOCUS Sequence 49 from Patent WO0073469.
DEFINITION AX056405
ACCESSION AX056405
VERSION AX056405.1 GI:12229112
KEYWORDS
SOURCE Murinae gen. sp.
ORGANISM Murinae gen. sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae.
REFERENCE 1 (bases 1 to 2615)
AUTHORS Plowman,G.D., Martinez,R., Whyte,D. and Sudersanam,S.
TITLE Protein kinases
JOURNAL Patent: WO 0073469-A 49 07-DEC-2000;
Sugen, Inc. (US)
FEATURES Location/Qualifiers
source 1..2615
/organism="Murinae gen. sp."
/db_xref="taxon:39108"
BASE COUNT 742 a 580 c 585 g 708 t
ORIGIN
Query Match 59.2%; Score 559.2; DB 6; Length 2615;
Best Local Similarity 81.8%; Pred. NO. 1.1e-115;
Matches 658; Conservative 0; Mismatches 143; Indels 3; Gaps 1;
QY 135 gaaaatagcactaagagaaaatacgtatgtlgaagcaattaaaacatccaaatcttgtgaa 194
Db 1 GAAAATAGCCCTGCGGGAAATCCGTATGCTGAAG---TTGAAACACCCAAACCTCGTGAA 57
QY 195 cctcatcgaggtgttcaggagaaaaaggaataatgcatttagtttgaatactgtgatca 254
Db 58 CCTCATCGAGGTGTTCAGAGAAAAGAGAAAGATGCATCTAGTTTIGAGTACTGTGATCA 117
QY 255 tacacttttaaatgagctggaaagaaaccccaaatggagttgctgatggagtgaataag 314
Db 118 CACACTGTTAAACGAGCTGGAGAGAAACCCAAACGGAGTTTCTGATGGAGTGATAAAAG 177
QY 315 cgtattatggcaaacacttcaagctcttaatttctgtcatatacataactgtattcacag 374

Db	178	 TGTCATATGGCAAAACCCITCAAGCCCTTAAC	1819 bp Sequence 13 from Patent WO0123579.	DNA	PAT	30-APR-2001
QY	375	agataaaacctgaaaatattcttaataactaagcaagggaataatcaagatttgtgactt	AX107724	1819 bp	DNA	
Db	238	GGATGTAAACCTGAAACACATCCTAATAACCAAGCAAGGGATGATAAAGATTTGTGACTT	AX107724			
QY	435	cggtttgcacaaattctgattccaggagatgcctacacocgattatgtagctacgagatg	AX107724.1	GI:13923203		
Db	298	TGGATTGTCAGGAATTCTTAATTCAGGAGACGCCCTACACAGACTATGTTGCCACCAAGGTG				
QY	495	gtaccgagctcctgaacttcttctgtggagagatactcagtatggttcttcaagtcgatatatg				
Db	358	GTACCGAGCCCCGAACCTCTCGTGGGAGACACGAAGTACGGTTCCCTGTAGACGTGTG				
QY	555	ggctattgggtgtgttttgcagagctcctgcagcagggccagggccacactgtggcctggaaaaatc				
Db	418	GGCCGTCCGGTGTGTTTTCAGAGCTCCTGACGGGTTCAGCCACTCTGGCCGGGAAAAATC				
QY	615	agatgtggaccaactttatctgataatcagaaacactaggaaaaattaatcccaagacatca				
Db	478	CGACGTGGACCAAGCTTTACCTGATCATCAGGACGTTGGGGAAGCTGATTCACAGACACCA				
QY	675	atcaatctttaaaagtaacgggttttccatggcatcagtatacctgttcccaactcctggagcagaagacat				
Db	538	GTCTATCTTTAGGAGTAACCAAGTTTTCGCCGGCATCAGCATACCTGAACCAAGGACAT				
QY	735	ggaactcttgaggaaaagtctcagatgttcatcctgtggctctgaacttcatgaaggg				
Db	598	GGAGACTCTTTGAAGAAAAATTCCAAATGTTCAGCCTGTGGCTTTAAGTTTCATCAAGGG				
QY	795	gtgtctgaagatgaatccagatgacagattaacactgttcccaactcctggagagctccta				
Db	658	ATGCCTGAAGATGAATCCTGATGAGAGGCTGACCTGTGCCAGCTGCTGGACAGTGCCCTA				
QY	855	ctttgattcttttbaagagggcccaattaaagaaagcacgcgtaatgaaggaaagaaacag				
Db	718	CTTTGAGTCTTTTCAAGAGGATCAAATGAAAAGAAAGCCCGCAGTCAGGGGAGAGCCG				
QY	915	aagacgccaacagggtacttccgct	938			
Db	778	AAGGCCCAGCAGAAATCAACTGCT	801			
RESULT	8					
LOCUS	AX107724					
DEFINITION	Sequence 13 from Patent WO0123579.					
ACCESSION	AX107724					
VERSION	AX107724.1					
KEYWORDS	human.					
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
AUTHORS	Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.					
TITLE	Human kinase proteins and polynucleotides encoding the same					
JOURNAL	Patent: WO 0123579-A 13 05-APR-2001;					
FEATURES	Lexicon Genetics Incorporated (US) Location/Qualifiers 1..1819 /organism="Homo sapiens" /db_xref="taxon:9606"					
BASE COUNT	539 a	368 c	406 g	505 t	1 others	
ORIGIN						

Query Match 53.7%; Score 507; DB 6; Length 1819;
Best Local Similarity 69.8%; Pred. No. 5.8e-104;
Matches 929; Conservative 0; Mismatches 0; Indels 402; Gaps 2;

QY	1	atggaaaagtatgaaaaattagctaagactggagaagggtcttattgggttgattcaaaa	60
Db	138	ATGGAAAAGTATGAAAAATTAGCTAAGACTGGAGAAGGGTCTTATGGGTTGTATTCAAA	197
QY	61	tcagaaaaacacctctctggacaagtagtagctgtttaaaaaatttctggaatctgaagat	120
Db	198	TcAGAAAACAAAACCTCTGGACAAC TAGTAGCTGTGTTAAAAAATTTGTGGAATCTGAAGAT	257
QY	121	gatacctgttgttaagaaaaatagcactaagagaataacgtatgttgaagcaattaaacat	180
Db	258	GATCCTGTGTTTAAGAAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACACAT	317
QY	181	ccaaatcttctgaacctcatcgaggtgttccagagagaaaaatgcatttagttttt	240
Db	318	CCAAATCTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAGGAAAATGCATTTAGTTTTT	377
QY	241	gaatactgtgatcatcacacttttaaatgagctgtggaagaaacccccaaatggagtgtgat	300
Db	378	GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCCAATGGAGTTGCTGAT	437
QY	301	ggagtatcaaaaagcgtattatggcaaacacttcaagctcttaatttctgtcatatacat	360
Db	438	GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCATATACAT	497
QY	361	aactgtattcacagagatatataaacctgaaaaatattctaataaactaagcaaggaataatc	420
Db	498	AACTGTATTACAGAGATATAAAACCTGAAAATATTCTAAATAACTAAGCAAGGAATAATC	557
QY	421	aagatttgtgacttcgggtttgtgcacaaaattctg	453
Db	558	AAGATTGTGACITCGGGTTTGCACAAATTCAGTGGACTTCATCTTCTCTGTGGTGC	617
QY	454	-----	453
Db	618	TCCTTGATTGGCTTAATAGTTGACCTTCTGAAATTCCTTTTCTGCCAATTCAGAGATTTT	677
QY	454	-----	453
Db	678	CTCCTGGCTTGGATCCATTGCTGACACAGTGTTTCACCATTGGGGCCAGGCTCATCTCGA	737
QY	454	-----	453
Db	738	ACTTCTGGCCTCAAGTGATCCTTCCACCTCGGCCCTCCCAAGTGTGGATTGCAAGTGTG	797
QY	454	-----	453
Db	798	AGCACCGTGGCCAGCCAGATTTTTCAAAACAATAACTACTGAGAGCTCACAAAGATTGTTT	857
QY	454	-----	474
Db	858	TTAGTGGGAACACAAATTCGAAACAAATTCCTTGAGAACGCCATCCAGGAGATGCCCTACACC	917
QY	475	gattatgtactacgagatggtacagagctcctgaacttcttgtgggagatactcagtat	534
Db	918	GATTATGTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTGTGGGAGATACTCAGTAT	977
QY	535	ggttcttcagtcgatatatgggclattgggtgtgttttgcagagctcctgacagggccag	594
Db	978	GGTTCTTCAGTCGATATATGGCTATTGGTTGTGTTTTCGAGAGCTCCTGACAGGCCAG	1037
QY	595	ccactgtggcctggaataatcagatgtggaccaactttatctgataatcagaacact----	650
Db	1038	CCACTGTGGCCTGGAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGTA	1097
QY	651	-----	650
Db	1098	GAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTCTTGACGTCAAGTGATCCA	1157
QY	651	-----	678
Db	1158	CCTGCCCTAGCCTCTCAAAGTGTGGAATTCAGAGGAAAAATTAATCCCAAGACATCAATCA	1217
QY	679	atctttaaaagtaaacgggtttttccatggcatcatcacctgagccagagacatgga	738

Matches		508;	Conservative	0;	Mismatches	352;	Indels	3;	Gaps	1;
QY	1	atggaaaaagtatgaaaaaattagctaagactggagaagggtcttatggggttgattcaaa								
Db	376	ATGGAAAAATATGAAAACCTGGGTTGGTTGGAGAAGGGAGTTATGGAATGGTGATCAAG								
QY	61	tgcagaaaaaaacctctggacaagtagtagctgttataaaaaatttggaaatctgaagat								
Db	436	TGTAGGAATAAGATAGCTGGAAGAAATGTGGCCATAAAGAAGTTCTTAGAAAGTCACGAT								
QY	121	gatcctgtgtttaagaaaaatagcaactaagagaaatacgtatgttgaaagcaattaaaaacat								
Db	496	GACAAAAATGGTTAAAAAGATTGCCAATGCCGAGAAATCAAGTTACTAAAGCAACTAGGCAT								
QY	181	ccaaatcttgtgaacctcatcgaggtgttcaggagaaaaaggaaaaatgcatttagttttt								
Db	556	GAAAACCTGGTGAATCTCTTGGAGTGTGTAAGAAAAAAGGATGGTACCTAGTCTTT								
QY	241	gaatactgtgatcatcacactttttaaagtgcgtggaaagaaaccccaaatggagtgcgat								
Db	616	GAATTTGTGACCCACACAATTTCTTGATGACTTGGAGCTCTTTCCAAATGGACTAGACTAC								
QY	301	ggagtgatcaaaagcgtattatggcaaacacttcaagctctttaatttctgtcatatacat								
Db	676	CAAGTAGTTCAAAAAGTATTCTTTCAGATTATTAAATGGAATTGGGAATTTGTACAGTCCAC								
QY	361	aactgtattcacagagatataaaaacctgaaaaattcttaataactaagcaaggaataatc								
Db	736	AATATCATACACAGAGATATAAAGCCAGAGAAATATATTAGTCTCCAGTCTGGCGTTGTC								
QY	421	aagatttgtgacttcgggtttgcacaaattctg---attccaggagatgcctacacccgat								
Db	796	AAGCTATCGGATTTTGGATTTGCGCGAACAATTGGCAGCTCTCTGGGAGGTTTATACTGAT								
QY	478	tatgtagctaogagatggtaccgagctcctgaaacttcttctgtgggagatactcagtatggt								
Db	856	TATGTGGCAACCCGATGGTACAGAGCTCCAGAACTATTGGTTGGTGATGTCAAGTATGGC								
QY	538	tcttcagtcgatatatatgggctattgtgtgttttgcagagctcctgcagggccagcca								
Db	916	AAGCTGTTGATGTGTGGGCCAATTGGTTGTCTGTAACCTGAAATGTTCAATGGGGGAACCC								
QY	598	ctgtggccctggaaaaatcagatgtggaccactttatctgataaatcagaacactaggaaaa								
Db	976	CTATTTCCCTGGAGATTTCTGATATTGATCAGCTATATCATATATGATGTGTTAGTAAAT								
QY	658	ttaatcccaagacatcaatcaaatctttaaaagtaacgggttttccatggcatcagtata								
Db	1036	CTAATTTCCAAGGCATCAGGAGCTTTTAAATAAAAATCCTGTGTTTGTGTGGAGTAAGGTTG								
QY	718	cctgagccagaagacatggaacctcttgaggaaaaagttctcagatgttcacacctgtggct								
Db	1096	CCTGAAATCAAGGAAGAGAACCTCTTGAAAGACCGCTATCCTAAGCTCTCTGAAGTGGTG								
QY	778	ctgaacttcatgaagggtgtctctgaagatgaatccagatgacagattaacctgttcccaa								
Db	1156	ATAGATTTAGCAAAAGAAATGCTTACATATTGACCCCGACAAAGACCCCTTCTGTGCTGAG								
QY	838	ctcctggagagctcctactttga								
Db	1216	CTCCTACACCATGATTTCTTTCA								

RESULT 14
AB029067
LOCUS
DEFINITION
AB029067 1518 bp mRNA
Mus musculus Kkm mRNA for Ser/Thr kinase KKIAMRE-delta, partial cds.
17-DEC-1999
ACCESSION
AB029067
VERSION
AB029067.1 GI:6594327
KEYWORDS
Ser/Thr kinase KKIAMRE-delta.
SOURCE
Mus musculus (strain:C57BL/6) adult brain tissue_lib:Uni-ZAP cdna
to mRNA.

ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	1 (sites)
TITLE	Sassa,T., Gomi,H., Sun,W., Ikeda,T., Thompson,R.F. and Itohara,S. The Murine KKIAMRE gene: variants, dual promoters, expression and chromosomal localization
JOURNAL	Unpublished (1999)
REFERENCE	2 (bases 1 to 1518)
AUTHORS	Sassa,T., Gomi,H. and Itohara,S.
TITLE	Direct Submission
JOURNAL	Submitted (16-JUN-1999) to the DBJ/EMBL/GenBank databases. Takayuki Sassa, The Institute of Physical and Chemical Research (RIKEN), Brain Science Institute (BSI); 2-1 Hiroosawa, Wako, Saitama 351-0198, Japan (E-mail:tasasa@brain.riken.go.jp, Tel:81-48-467-9724(ex.7716), Fax:81-48-467-9725)
FEATURES	Location/Qualifiers
source	1..1518 /organism="Mus musculus" /strain="C57BL/6" /db_xref="taxon:10090" /dev_stage="adult" /tissue_lib="Uni-ZAP" /tissue_type="brain" 530..1518 /gene="Kkm" 530..>1518 /gene="Kkm" /note="cdc2-related kinase" /codon_start=1 /product="Ser/Thr kinase KKIAMRE-delta" /protein_id="BAA88429.1" /db_xref="GI:6594328"
gene	/translation="MEKYENLGLVGEISYGMVKCRNKDSGRIVAIKKFLESDDDKMV KKIAMREIKLLKQLRHENLVNLEVCCKKKRWYLVFEFVDHTILDILKLFNGLDYQV VQYLFQILINGIFGCHSHNIIHRDIKPENILVSQSVKLCDFGFARTLAAPGEVYTD YVATRWYRAPELLVGDVKYKAVDIWAIGCLVIEMLMGQPLFPGESIDQLHHIMTCL GNLIPRHQELFYKNPVFAGVRLPEVKDAEAPLESRYPKLPEAVISLAKKCLHIDPDK RPFCDLLRHDFDOMDGFARFISOELQLKTEKDARNNSLPKKSQKKKKKKKK"
CDS	BASE COUNT 421 a 303 c 382 g 412 t ORIGIN

Query Match	28.8%;	Score 272.6;	DB 10;	Length 1518;	
Best Local Similarity	58.8%;	Pred. No. 2.7e-51;			
Matches 511;	Conservative	0;	Mismatches 349;	Indels 9; Gaps 2;	
QY	1	atggaaaaagtatgaaaaaattagctaagactggagaagggtctttatggggttgattcaaa			60
Db	530	ATGGAGAAGTACGAGAACCTAGGATTGGTTGGAGAGGGGAGTTACGGAATGGTGATGAAG			589
QY	61	tgcagaaaaaaacctctggacaagtagtagctgttaaaaaatttggaaatctgaagat			120
Db	590	TGTAGGAACAAAGACAGTGGAGAAATGTGGCCATTAAAGAAAGTTCTTAGAAAGTGACGAT			649
QY	121	gatccgttgtttaagaaaaatagcactaagagaaaaatcgtatgttgaagcaattaaaaacat			180
Db	650	GATAAAATGGTTAAGAAGATCGCAATCGGAGAGATCAAGTTACTAAAGCAACTGCGGCAT			709
QY	181	ccaaatcttgtgaacctcatcgagggtgttcaggagaaaaaagaaatgcatttagtttt			240
Db	710	GAAAATTTGGTGAATCTGTTGGAAGTAGTAAGAAAAAAAACGTTGGTACCTAGTCTTT			769
QY	241	gaatactgtgatcatcacactttttaaatagctggagaaaaacccaaatggagttgctgat			300
Db	770	GAAATTTGTGACCACACGATCTTTGATGACTTGAAACTCTTTCCAAATGGACTAGACTAT			829
QY	301	ggagtgatcaaaagcgtattatggcaaacacttcaagctcttaatttctgtcatatacat			360
Db	830	CAAGTAGTTCAAAAATATTGTTTCAGATTATTAAATGGAATTTGGATTTGCCACAGTCAC			889
QY	361	aactgtattcacagagatataaaacctgaaaaatcttaataactaagcaagggaataatc			420

Db	890	AAATATCATACACAGAGATATAAAGCCTGAGAATATATTAGTCTCCCACTCTGGCGTTGTA	949
QY	421	aagatttgtgacttcgggttgccacaaattct---gattccaggagatgcctacaccgat	477
Db	950	AAATTATGTGACTTTGGATTTCACGGACCCCTGGCGCTCCAGGGAGGTTTACACTGAC	1009
QY	478	tatgtagctacgagatggtacogagctcctgaacttcttgtgggagatactcagtatggt	537
Db	1010	TATGTGCCCACCCGGTGTACAGAGCTCCAGAACTGCTGGTGGTGACGTTCAAGTATGGC	1069
QY	538	tcttcagtcgatatatgggcttattggttggttttttcagagctcctgacagggccagcca	597
Db	1070	AAGGCTGTGATATCTGGGCCATGGTTGTCTGTCATTGAAATGCTTATGGGCGAGCCA	1129
QY	598	ctgtggcctggaaaaatcagatgtggaccacttttatctgataatacagaacactaggaaaa	657
Db	1130	CTGTTTCCAGGAGAACTGATATTGATCAGCTACATCATATTATGACATGTTTAGGTAAT	1189
QY	658	ttaatcccaagacatcaatcaatttttaaaagtaacgggtttttcccatggcatcagtata	717
Db	1190	TTAATTCCAAGACACCAGGAGCTATTTATAAAATCCTGTGTTGCTGGAGTAAGGTTG	1249
QY	718	cctgagccagaagacatggaa-----actcttgaggaaaagttctcagatgttcatcct	771
Db	1250	CCTGAGGTCAAAGATGCAGAACGACAGCCTCTTGAGAGCCGCTATCCTAAGCTCCCTGAA	1309
QY	772	gtggctctgaacttcaatgaaggggtgctgaagatgaatccagatgacagattaaacctgt	831
Db	1310	GCTGTGATCAGTTTAGCAAGAAATGCTTGCACATTGACCCAGACAAAAAGGCCCTTCTGT	1369
QY	832	tcccaactcctggagagctcctcactttga	860
Db	1370	GCTGACCTTCTGCGCCATGACTTCTTTCA	1398
RESULT	15		
AB029066			
LOCUS	AB029066	3988 bp	mRNA
DEFINITION	Mus musculus Kkm mRNA for Ser/Thr kinase KKIAMRE-gamma, complete cds.	ROD	17-DEC-1999
ACCESSION	AB029066		
VERSION	AB029066.1	GI:6594325	
KEYWORDS	Ser/Thr kinase KKIAMRE-gamma.		
SOURCE	Mus musculus (strain:C57BL/6) adult brain tissue	lib:Uni-ZAP	cDNA
ORGANISM	Mus musculus		
REFERENCE	1 (sites)		
AUTHORS	Sassa,T., Gomi,H., Sun,W., Ikeda,T., Thompson,R.F. and Itohara,S.		
TITLE	The Murine KKIAMRE gene: variants, dual promoters, expression and chromosomal localization		
JOURNAL	Unpublished (1999)		
REFERENCE	2 (bases 1 to 3988)		
AUTHORS	Sassa,T., Gomi,H. and Itohara,S.		
TITLE	Submitted (16-JUN-1999) to the DDBJ/EMBL/GenBank databases.		
JOURNAL	Takayuki Sassa, The Institute of Physical and Chemical Research (RIKEN), Brain Science Institute (BSI); 2-1 Hirosawa, Wako, Saitama 351-0198, Japan (E-mail:tasasa@brain.riken.go.jp, Tel:81-48-467-9724(ex.7716), Fax:81-48-467-9725)		
FEATURES	Location/Qualifiers		
source	1. .3988		
	/organism="Mus musculus"		
	/strain="C57BL/6"		
	/db_xref="taxon:10090"		
	/dev_stage="adult"		
	/tissue_lib="Uni-Zap"		
	/tissue_type="brain"		
	240. .1934		
gene	/gene="Kkm"		
CDS	240. .1934		
	/gene="Kkm"		

		BASE COUNT	1238 a	733 c	901 g	1116 t	
		ORIGIN					
		Query Match		28.8%;	Score 272.6;	DB 10;	Length 3988;
		Best Local Similarity		58.8%;	Pred. No. 2.7e-51;		
		Matches 511;	Conservative	0;	Mismatches 349;	Indels	9; Gaps 2;
QY	1	atggaagaagtataaaaaattagctaagactggagaagggtcttattatgggttgattcaaa	60				
Db	240	ATGGAGAAGTACGAGAACCTAGGATTGGTTGGAGAGGGGAGTTACGGAATGGTGATGAAG	299				
QY	61	tcgagaacaaaaacctctctggacaagtagtagctgttaaaaaatttgtggaatctgaaagt	120				
Db	300	TGTAGGAACAAAAGACAGCTGGAAGAAATTGTGGCCATTAAAGAAAGTTCTTAGAAAGTGACGAT	359				
QY	121	gatactgtgtttaagaaaaatagcacactaagagaaaaatacgtatgttgaagcaaatataaacat	180				
Db	360	GATAAAATGGTTAAGAAGATCGCAATGCGAGAGATCAAGTTACTAAAGCAACTGCGGCAT	419				
QY	181	ccaaatcttgtgaacctcatcgagggtgttcaggagagaaaaaggaatgcatttagtttt	240				
Db	420	GAAATTTGGTGAATCTGTTGGAAGTATCTAAGAAAAAAGAAACGTTGGTACCTAGTCTTT	479				
QY	241	gaatactgtgatcatcacacttttaaatgagctggaaagaaacccaaatggagttgctgat	300				
Db	480	GAA'TTGTTGACCCACACGATTCTTGATGACTTGAAC'TCTTTCCAATGGACTAGACTAT	539				
QY	301	ggagtcatcaaaaagcgtattatggcaaacacttcaagctcttaatttctgtcatatacat	360				
Db	540	CAAGTAGTTCAAAAATATTTGTTTCAGATTATTAATGGAATGGATT'TTGGCCACAGTCA	599				
QY	361	aactgtattcacagagatatataaaccttgaaaaatattctataataactaagcaaggaataatc	420				
Db	600	AATATCATACACAGAGATATAAAGCCTGAGATATATTACTCTCCACAGCTCTGGCGTTGTA	659				
QY	421	aagatttgcacttcgggttgcacaaattct---gattccaggagatgcctacaccgat	477				
Db	660	AAATTATGTGACTTTGGATTTCACGGACCCCTGGCGGCTCCAGGGGAGGTTTACACTGAC	719				
QY	478	tatgtagctacgagatggtaccgagctcctgaacttcttgtgggagatactcagtatggt	537				
Db	720	TATGTGCCCACCCGGTGGTACAGAGCTCCAGAACTGCTGGTGGGTCAAGTATGGC	779				
QY	538	tcttcagtcgatatatgggctattggttgtgttttttgcagagctcctgacagggccagcca	597				
Db	780	AAGGCTGTGGATATCTGGGCCATTGGTTGTCGTCTGTCATTGAAATGCTTATGGGGCAGCCA	839				
QY	598	ctgtggcctggaaaaatcagatgtggacaaactttatcttgataatacagaacactaggaaaa	657				
Db	840	CTGTTTCCAGGAGAACTCTGATATTGATCAGCTACATCATATTATGACATGTTTAGGTAAT	899				
QY	658	ttaatcccaagacatcaatcattttaaaagtacgggtttttccatggcatcagtata	717				
Db	900	TTAATTCCAAGACACCAGGAGCTATTTTATAAAATCCTGTGTGCTGGAGTAAGGTTG	959				
QY	718	cctgagccagaagacatggaa-----actcttgagaaaaagttctcagatgttcatcct	771				

/note="cdc2-related kinase"
/codon_start=1
/product="Ser/Thr kinase KKIAMRE-gamma"
/protein_id="BAA88428.1"
/db_xref="GI:6594326"
/translation="MEKYENLGLVGEYSYGMVMKCRNKDSGRIVAIAKKFLESDDDKMY
KKIAMREIKLLQLRHENLVNLEVCCKKKRWYLVFEFVDHTILDDLKLPNGLDYQV
VQYLFQIINGIGFCHSHNIHRDIKENILVSQSGVVKLCDFGFARTLAAPGEVYTD
YVATRWYRAPELLVDVKYKAVDIWAIGCLVIEMLMGQPLFPGESDIDQLHHMTCL
GNLIPRHOELFYKNPVFAGVRLPEVKDAEAEPLERYPKLPEAVISLAKKLHIDPK
RPFADLLRHDFQMDGFAERFSOEQLKIEKDARNNSLPKKSQRKKEKDDALGEER
KTLVVQDTNADPKIKDSKVFVKVSGKIDVEKMEKSGRSNANCLHDNGTNHKGLASTS
LRDCSNVNIDHSRNPGTAIPEPLTHNLSAVAPGINAGMTIPGVQNYRVDEKTKKYNP
FVKPNQPPPPAGIYNMNVSTSVSGEKYLLQANKRKEYPKADVRLPELNYNHLPELRAL
EGIARNSRLIKKENKCLSESRIPSLAIDLHVSSVASHQMLPALEMTSSCTRALIP"

Db 960 CCTGAGGTCAAAGATGCAGAAGCAGAGCCTCTTGAGAGCCGGCTATCCTAAGCTCCCTGAA 1019
QY 772 gtggctctgaacttcatgaagggtgtgtctgtgaagatgaatccagatgacagattaaacctgt 831
Db 1020 GCTGTGATCAGTTTAGCAAAGAAATGCTTGACATTTGACCCAGACAAAAGGCCCTTCTGT 1079
QY 832 tcccaactcctggagagctcctactttga 860
Db 1080 GCTGACCTTCTGCGCCATGACTTCTTCA 1108

Search completed: May 2, 2002, 16:29:40
Job time: 9562 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 17:13:01 ; Search time 234.75 Seconds
(without alignments)
3451.213 Million cell updates/sec

Title: US-09-671-050-11
Perfect score: 945
Sequence: 1 atggaaaagtatgaaaatt.....aggctacttcgcgtcaaaagt 945

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :		N_Geneseq_1101.*
1:	/SIDS1/gcgdata/geneseq/geneseqn/NA1980.DAT.*	
2:	/SIDS1/gcgdata/geneseq/geneseqn/NA1981.DAT.*	
3:	/SIDS1/gcgdata/geneseq/geneseqn/NA1982.DAT.*	
4:	/SIDS1/gcgdata/geneseq/geneseqn/NA1983.DAT.*	
5:	/SIDS1/gcgdata/geneseq/geneseqn/NA1984.DAT.*	
6:	/SIDS1/gcgdata/geneseq/geneseqn/NA1985.DAT.*	
7:	/SIDS1/gcgdata/geneseq/geneseqn/NA1986.DAT.*	
8:	/SIDS1/gcgdata/geneseq/geneseqn/NA1987.DAT.*	
9:	/SIDS1/gcgdata/geneseq/geneseqn/NA1988.DAT.*	
10:	/SIDS1/gcgdata/geneseq/geneseqn/NA1989.DAT.*	
11:	/SIDS1/gcgdata/geneseq/geneseqn/NA1990.DAT.*	
12:	/SIDS1/gcgdata/geneseq/geneseqn/NA1991.DAT.*	
13:	/SIDS1/gcgdata/geneseq/geneseqn/NA1992.DAT.*	
14:	/SIDS1/gcgdata/geneseq/geneseqn/NA1993.DAT.*	
15:	/SIDS1/gcgdata/geneseq/geneseqn/NA1994.DAT.*	
16:	/SIDS1/gcgdata/geneseq/geneseqn/NA1995.DAT.*	
17:	/SIDS1/gcgdata/geneseq/geneseqn/NA1996.DAT.*	
18:	/SIDS1/gcgdata/geneseq/geneseqn/NA1997.DAT.*	
19:	/SIDS1/gcgdata/geneseq/geneseqn/NA1998.DAT.*	
20:	/SIDS1/gcgdata/geneseq/geneseqn/NA1999.DAT.*	
21:	/SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT.*	
22:	/SIDS1/gcgdata/geneseq/geneseqn/NA2001.DAT.*	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	945	100.0	945	22	Human kinase cDNA
2	929	98.3	972	22	Human kinase cDNA
3	863	91.3	1083	22	Polynucleotide seq
4	839	88.8	1041	22	Human kinase cDNA
5	823	87.1	1068	22	Human kinase cDNA
6	643	68.0	911	22	Novel protein kina
7	559.2	59.2	2615	22	Novel protein kina
8	507	53.7	1819	22	Human kinase cDNA
9	454.4	48.1	561	22	Human kinase cDNA
10	454.4	48.1	594	22	Human kinase cDNA
11	287	30.4	1698	21	Rabbit KKIAMRE kin

12	287	30.4	3080	21	AAA29745	Rabbit KKIAMRE kin
13	227.4	24.1	1513	21	AAZ51208	Human lost in leuk
14	227.4	24.1	1773	22	AAS06724	Polynucleotide seq
15	211.4	22.4	1667	21	AAZ51207	Rat lost in leukae
16	211.4	22.4	1738	21	AAZ51206	Rat lost in leukae
17	175.8	18.6	1866	22	AAF44672	Novel protein kina
18	171.8	18.2	903	20	AAZ07476	Pneumocystis carin
19	163.4	17.3	1476	21	AAZ37835	Cyclin-dependent k
20	163.4	17.3	1635	19	AAV71073	CDK2-green fluores
21	163.4	17.3	1635	19	AAV71074	Green fluorescent
22	161.4	17.1	1322	21	AAF15889	Human prostate can
23	136.6	14.5	1825	22	AAI58432	Human polynucleoti
24	134.6	14.2	1089	15	AAQ63490	Cyclin dependent p
25	131	13.9	1002	18	AAT64448	CDK1 protein. Can
26	131	13.9	1002	21	AAZ08264	Candida gene encod
27	128.2	13.6	6015	21	AAZ75931	Human ORFX ORF1486
28	128	13.5	1310	21	AAC35344	Arabidopsis thalia
29	126.6	13.4	1899	22	AAF44678	Novel protein kina
30	126.4	13.4	1307	21	AAC51531	Arabidopsis thalia
31	126	13.3	1293	21	AAC36413	Arabidopsis thalia
32	125.6	13.3	1788	22	AAT160218	Human polynucleoti
33	125.2	13.2	1326	21	AAZ43758	Zea mays DNA fragm
34	117.6	12.4	936	22	AAF58252	Oligonucleotide D1
35	117.6	12.4	936	22	AAF58254	Oligonucleotide D1
36	117.6	12.4	936	22	AAF58257	Oligonucleotide D1
37	117.6	12.4	936	22	AAF58259	Oligonucleotide D2
38	117.6	12.4	936	22	AAF58262	Oligonucleotide D2
39	117.6	12.4	938	22	AAF58255	Oligonucleotide D1
40	117.2	12.4	936	22	AAF58252	Oligonucleotide D1
41	117.2	12.4	936	22	AAF58254	Oligonucleotide D1
42	117.2	12.4	936	22	AAF58257	Oligonucleotide D1
43	117.2	12.4	936	22	AAF58259	Oligonucleotide D2
44	117.2	12.4	936	22	AAF58262	Oligonucleotide D2
45	117.2	12.4	938	22	AAF58255	Oligonucleotide D1

ALIGNMENTS

RESULT 1	
AAD03817	
ID	AAD03817 standard; cDNA; 945 BP.
XX	
AC	AAD03817;
DT	19-JUN-2001 (first entry)
XX	
DE	Human kinase cDNA #6.
XX	
KW	Human; kinase; gene therapy; bioreactor; mental disorder;
KW	biological disorder; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	1..945
FT	/*tag= a
FT	/product= "Human kinase #6"
FT	/note= "The coding region does not include stop codon"
FT	/partial
XX	
PN	WO200123579-A1.
XX	
PD	05-APR-2001.
XX	
PF	27-SEP-2000; 2000WO-US26621.
XX	
PR	28-SEP-1999; 99US-0156511.
XX	
PA	(LEXI-) LEXICON GENETICS INC.
XX	
PI	Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
XX	

DR WPI; 2001-266166/27.
XX P-PSDB; AAE00495.
PT New isolated human kinase polynucleotide useful for generating
PT antibodies, as reagents in diagnostic assays and for screening for
PT compounds useful for treating mental, biological or medical diseases
XX
PS Claim 3; Page 33; 38pp; English.
XX
CC The present sequence is a cDNA encoding novel human protein
CC (NHP) known as human kinase. The human kinases share structural
CC similarity with animal kinases, more particularly serine or
CC threonine protein kinases. Human kinase cDNA is useful for the
CC detection of mutant human kinase for the diagnosis of disease,
CC and also as a therapeutic. It is useful for screening drugs
CC effective in the treatment of symptomatic or phenotypic
CC manifestations perturbing the normal function of NHP in the
CC body. The NHP nucleotide sequences are useful for generation of
CC antibodies, as reagents in diagnostic assays, for the
CC identification of other cellular gene products related to human
CC kinases, and as reagents in assays for screening compounds that
CC are useful for treating mental, biological or medical disorders.
CC NHP oligonucleotides are used as probes. The labelled NHP probes
CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.
XX
SQ Sequence 945 BP; 323 A; 165 C; 202 G; 255 T; 0 other;

Query Match 100.0%; Score 945; DB 22; Length 945;
Best Local Similarity 100.0%; Pred. No. 4.4e-245;
Matches 945; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atggaaaagtatgaaaaattagctaagactggagaagggtcttatggggtgtattcaaa 60
Db |||||
QY 1 atggaaaagtatgaaaaattagctaagactggagaagggtcttatggggtgtattcaaa 60
Db |||||
QY 61 tgcagaaaacaaacctctggacaagtagtagctgttaaaaaatttggaaatctgaagat 120
Db |||||
QY 61 tgcagaaaacaaacctctggacaagtagtagctgttaaaaaatttggaaatctgaagat 120
Db |||||
QY 121 gatcctgtgttaagaaaaatagcactaagagaaatacgtatgttgaagcaattaaaaacat 180
Db |||||
QY 121 gatcctgtgttaagaaaaatagcactaagagaaatacgtatgttgaagcaattaaaaacat 180
Db |||||
QY 181 ccaaatcttgtgaacctcatcgagggtgttcaggagaaaaaggaaatgcattagtttt 240
Db |||||
QY 181 ccaaatcttgtgaacctcatcgagggtgttcaggagaaaaaggaaatgcattagtttt 240
Db |||||
QY 241 gaatactgtgatcatacacttttaaatgagctggaagaaacccaaatggagttgctgat 300
Db |||||
QY 241 gaatactgtgatcatacacttttaaatgagctggaagaaacccaaatggagttgctgat 300
Db |||||
QY 301 ggagtgatcaaaagcgtattatggcaacacttcaagctcttaatttctgcataatacat 360
Db |||||
QY 301 ggagtgatcaaaagcgtattatggcaacacttcaagctcttaatttctgcataatacat 360
Db |||||
QY 361 aactgtattcacagagatataaaacctgaaataattctaataactaagcaaggaataatc 420
Db |||||
QY 361 aactgtattcacagagatataaaacctgaaataattctaataactaagcaaggaataatc 420
Db |||||
QY 421 aagatttgtgacttcgggtttgcacaaattctgattccaggagatgcctacaccgattat 480
Db |||||
QY 421 aagatttgtgacttcgggtttgcacaaattctgattccaggagatgcctacaccgattat 480
Db |||||
QY 481 gtagctacagatggtaccgagctcctgaaacttttgtgggagatactcagtagttct 540
Db |||||

Db 481 gtagctacagatggtaccgagctcctgaaacttcttgggagatactcagtagttct 540
QY 541 tcagtcgatatatgggctattgtgtgttttgcagagctcctgacagggccagccactg 600
Db |||||
QY 541 tcagtcgatatatgggctattgtgtgttttgcagagctcctgacagggccagccactg 600
Db |||||
QY 601 tggcctggaaaaatcagatgtggaccaactttatctgataatcagaacactaggaaaaatta 660
Db |||||
QY 601 tggcctggaaaaatcagatgtggaccaactttatctgataatcagaacactaggaaaaatta 660
Db |||||
QY 661 atcccaagacatcaatcaatctttaaaagttaacgggttttccatggcatcagataacct 720
Db |||||
QY 661 atcccaagacatcaatcaatctttaaaagttaacgggttttccatggcatcagataacct 720
Db |||||
QY 721 gagccagaagacatggaaaactcttgagaaaaagttctcagatgttcatcctgtggtctcg 780
Db |||||
QY 721 gagccagaagacatggaaaactcttgagaaaaagttctcagatgttcatcctgtggtctcg 780
Db |||||
QY 781 aacttcatgaagggtgtctgaagatgaatccagatgacagattaacctgttcccaactc 840
Db |||||
QY 781 aacttcatgaagggtgtctgaagatgaatccagatgacagattaacctgttcccaactc 840
Db |||||
QY 841 ctggagagctcctactttgatcttttcaagagagcccaaatataaagaaagcacgtaat 900
Db |||||
QY 841 ctggagagctcctactttgatcttttcaagagagcccaaatataaagaaagcacgtaat 900
Db |||||
QY 901 gaaggaagaaacagagaagacgcccacaggttacttccgctcaaaaagt 945
Db |||||
QY 901 gaaggaagaaacagagaagacgcccacaggttacttccgctcaaaaagt 945
Db |||||

RESULT 2
AAD03814
ID AAD03814 standard; cDNA; 972 BP.
AC AAD03814;
XX
DT 19-JUN-2001 (first entry)
XX Human kinase cDNA #3.
DE Human; kinase; gene therapy; bioreactor; mental disorder;
XX biological disorder; ss.
KW Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..972
FT /*tag= a
FT /product= "Human kinase #3"
FT /note= "The coding region does not include stop codon"
FT /partial
XX
PN WO200123579-A1.
XX
XX 05-APR-2001.
XX
XX 27-SEP-2000; 2000WO-US26621.
XX
XX 28-SEP-1999; 99US-0156511.
XX
XX (LEXI-) LEXICON GENETICS INC.
XX
XX Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
PI WPI; 2001-266166/27.
XX P-PSDB; AAE00492.
DR
XX
PT New isolated human kinase polynucleotide useful for generating
PT antibodies, as reagents in diagnostic assays and for screening for
PT compounds useful for treating mental, biological or medical diseases
XX

PS	Disclosure; Page 29-30; 38pp; English.
XX	
CC	The present sequence is a cDNA encoding novel human protein
CC	(NHP) known as human kinase. The human kinases share structural
CC	similarity with animal kinases, more particularly serine or
CC	threonine protein kinases. Human kinase cDNA is useful for the
CC	detection of mutant human kinase for the diagnosis of disease,
CC	and also as a therapeutic. It is useful for screening drugs
CC	effective in the treatment of symptomatic or phenotypic
CC	manifestations perturbing the normal function of NHP in the
CC	body. The NHP nucleotide sequences are useful for generation of
CC	antibodies, as reagents in diagnostic assays, for the
CC	identification of other cellular gene products related to human
CC	kinases, and as reagents in assays for screening compounds that
CC	are useful for treating mental, biological or medical disorders.
CC	NHP oligonucleotides are used as probes. The labelled NHP probes
CC	are useful for screening human genomic library for identifying
CC	polymorphisms and as primers in amplification assays to detect
CC	mutations within the exons, introns and splice sites that can
CC	be used in diagnostics and pharmacogenomics. Nucleotide construct
CC	encoding NHP products are used to genetically engineer cells
CC	in vivo that functions as bioreactors in the body delivering a
CC	continuous supply of NHP to the body. Nucleotide constructs
CC	encoding functional NHPs are used in gene therapy for the
CC	modulation of NHP expression.
XX	
SQ	Sequence 972 BP; 323 A; 172 C; 214 G; 263 T; 0 other;
Query Match 98.3%; Score 929; DB 22; Length 972;	
Best Local Similarity 100.0%; Pred. No. 9.1e-241;	
Matches 929; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 atggaaaagtatgaaaaattagctaagactggagaagggtcttatgggtgtattcaaa 60
Db	
QY	1 atggaaaagtatgaaaaattagctaagactggagaagggtcttatgggtgtattcaaa 60
Db	
QY	61 tgcagaaacaaacctctggacaagttagctgttaaaaaatttggaaatctgaagat 120
Db	
QY	61 tgcagaaacaaacctctggacaagttagctgttaaaaaatttggaaatctgaagat 120
Db	
QY	121 gatcctgtgttaagaaaaatagcactaagagaaaatcgtatgtgaagcaattaaaaacat 180
Db	
QY	121 gatcctgtgttaagaaaaatagcactaagagaaaatcgtatgtgaagcaattaaaaacat 180
Db	
QY	181 ccaaatcttgtgaacctcgcaggtgttcaggagaaaaaggaaaatgcatttagtttt 240
Db	
QY	181 ccaaatcttgtgaacctcgcaggtgttcaggagaaaaaggaaaatgcatttagtttt 240
Db	
QY	241 gaatactgtgatcacacttttaaatgagctggaaagaaaccccaatggagtgtgat 300
Db	
QY	241 gaatactgtgatcacacttttaaatgagctggaaagaaaccccaatggagtgtgat 300
Db	
QY	301 ggagtgatcaaaagcgtattatggcaaacacttcaagctcttaattctgtcatatacat 360
Db	
QY	301 ggagtgatcaaaagcgtattatggcaaacacttcaagctcttaattctgtcatatacat 360
Db	
QY	361 aactgtattcacagagatataaaacctgaaaatattcttaataactaagcaaggaataatc 420
Db	
QY	361 aactgtattcacagagatataaaacctgaaaatattcttaataactaagcaaggaataatc 420
Db	
QY	421 aagatttggacttcgggtttgcacaaaattctgattccaggagatgcctacaccgattat 480
Db	
QY	421 aagatttggacttcgggtttgcacaaaattctgattccaggagatgcctacaccgattat 480
Db	
QY	481 gtagctacgagatggtaccgagctctgaaacttcttggaggagatactcagtatggttct 540
Db	
QY	481 gtagctacgagatggtaccgagctctgaaacttcttggaggagatactcagtatggttct 540
Db	
QY	541 tcagtcgatataatgggcatttgggtgtgttttttcagagagctcctgacagccagccactg 600
Db	
QY	541 tcagtcgatataatgggcatttgggtgtgttttttcagagagctcctgacagccagccactg 600
Db	

QY	601 tggcctggaaaaatcagatgtggaccacacttttatctgataataatcagaacactaggaaaaatta 660
Db	
QY	601 tggcctggaaaaatcagatgtggaccacacttttatctgataataatcagaacactaggaaaaatta 660
Db	
QY	661 atccaagacatcaaatcaatctttaaaagaacgggttttccatggcatcagtataacct 720
Db	
QY	661 atccaagacatcaaatcaatctttaaaagaacgggttttccatggcatcagtataacct 720
Db	
QY	721 gagccagaagacatggaaaactcttgaggaaaagtctcagatgttccatcctgtggtctctg 780
Db	
QY	721 gagccagaagacatggaaaactcttgaggaaaagtctcagatgttccatcctgtggtctctg 780
Db	
QY	781 aacttcataaggggtgtctgaagatgaatccagatgacagattaacctgttcccaactc 840
Db	
QY	781 aacttcataaggggtgtctgaagatgaatccagatgacagattaacctgttcccaactc 840
Db	
QY	841 ctggagagctcctacttcttattctttcaagagggcccaattaaaagaaaagcagtaat 900
Db	
QY	841 ctggagagctcctacttcttattctttcaagagggcccaattaaaagaaaagcagtaat 900
Db	
QY	901 gaaggaagaacacagaaagacgcgccaacaggt 929
Db	
QY	901 gaaggaagaacacagaaagacgcgccaacaggt 929
Db	
RESULT 3	
AAS06725	
ID	AAS06725 standard; cDNA; 1083 BP.
XX	
AC	AAS06725;
XX	
DT	12-SEP-2001 (first entry)
XX	
DE	Polynucleotide sequence encoding human protein kinase #25.
XX	
KW	Human; protein kinase; PRK; STK; cancer; cardiovascular disease;
KW	metabolic disorder; immune related disease; neurological disorder;
KW	neurodegenerative disorder; inflammatory disorder; infectious disease;
KW	reproductive disorder; gene therapy; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200138503-A2.
XX	
PD	31-MAY-2001.
XX	
PF	22-NOV-2000; 2000WO-US32085.
XX	
PR	24-NOV-1999; 99US-0167482.
XX	
PA	(SUGE-) SUGEN INC.
XX	
PI	Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
PI	Flanagan P, Clary D;
XX	
DR	WPI; 2001-343950/36.
DR	P-PSDB; AAU03525.
XX	
PT	Nucleic acids encoding human kinase polypeptides, useful for preventing
PT	diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT	neuronal-associated diseases, and microbial infections -
XX	
PS	Example 1; Figure 1; 433pp; English.
XX	
CC	AAS06701-AAS06757 encode for novel human protein kinases #1-57. The
CC	novel protein kinases have been identified as members of the tyrosine
CC	or serine/threonine kinase (PRK and STK) families. The polynucleotides
CC	encoding protein kinases and the polypeptides may be used in the
CC	prevention, diagnosis and treatment of diseases associated with
CC	inappropriate kinase expression. For example, they may be used to treat
CC	cancers (especially cancers of haematopoietic origin), cardiovascular
CC	disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC	immune related diseases (e.g. rheumatoid arthritis), neurological

disorders (e.g. schizophrenia), neurodegenerative disorders (e.g. parkinson's disease), inflammatory disorders (e.g. asthma), infectious disease (e.g. HIV) and reproductive disorders (e.g. infertility). Additionally, polynucleotides encoding protein kinases may be used for gene therapy and as DNA probes in diagnostic assays. The protein kinase polypeptides may be used as antigens in the production of antibodies against the protein kinases and in assays to identify modulators of protein kinase expression and activity.

```

Query Match          91.3%; Score 863; DB 22; Length 1083;
Best Local Similarity 94.5%; Pred. NO. 5.8e-223;
Matches 927; Conservative 0; Mismatches 0; Indels 54

```

QY	1	atggaagaagtatgaaaaattagctaaagactggagaaggtcttatgagggttgatttcaaaa	60
Db	1		
QY	61	tgcagaacacaaacctctggacaagttagctgtttaaaaaatttgtggaatctgaagat	120
Db	61		
QY	121	gatactgttgttaagagaaatagcactaagagaaatacgtatgttgaagcaattaaaaaat	180
Db	121		
QY	181	ccaaatcttgtgaacctcatcgaggtgttcaggagaaaaaggaaaaatgcatttagtttt	240
Db	181		
QY	241	gaatactgtgatacatcaccttttaaatgagctggaaagaaacccaaaatggagttgtgat	300
Db	241		
QY	301	ggagtgatcaaaagcgtattatggcaaacacttcaagctctttaatttctgtcatacatat	360
Db	301		
QY	361	aactgtattcacagagataaaaaoctgaaaataattctataactaagcaaggaataatc	420
Db	361		
QY	421	aagatttggacttcgggttggacaaattctga-----	454
Db	421		
QY	455	-----	
Db	481	ttacctgatctgatgatgcatttgcaggttccaggagatgcctacaccgattatgtagct	540
QY	487	acgagatggtaccgagctcctgaacttcttgtgggagatactcagtatggttcttcagtc	546
Db	541	acgagatggtaccgagctcctgaacttcttgtgggagatactcagtatggttcttcagtc	600
QY	547	gatatatgggctattggttgtgttttggcagagctcctgacaggccagccactgtggcct	606
Db	601	gatatatgggctattggttgtgttttggcagagctcctgacaggccagccactgtggcct	660
QY	607	ggaaaatcagatgtggaccaacttttatctgataatcagaacactaggaaaaattaatccca	666
Db	661	ggaaaatcagatgtggaccaacttttatctgataatcagaacactaggaaaaattaatccca	720
QY	667	agacatcaatcaatctttaaaaagtaacgggttttccatggcatcagtaacacctgagcca	726
Db	721	agacatcaatcaatctttaaaaagtaacgggttttccatggcatcagtaacacctgagcca	780
QY	727	gaagacatggaactctttaggaaaaagtctctcagatgttctcactgtggtctgaacttc	786
Db	781	gaagacatggaactctttaggaaaaagtctctcagatgttctcactgtggtctgaacttc	840
QY	787	atgaagggtgtctgaagatgaatccagatgacagattaaacctgttcccaactcttgaag	846

CC	encoding NHP products are used to genetically engineer cells	
CC	in vivo that functions as bioreactors in the body delivering a	
CC	continuous supply of NHP to the body. Nucleotide constructs	
CC	encoding functional NHPs are used in gene therapy for the	
CC	modulation of NHP expression.	
XX		
SQ	Sequence 1041 BP; 344 A; 190 C; 228 G; 279 T; 0 other;	
	Query Match 88.8%; Score 839; DB 22; Length 1041;	
	Best Local Similarity 90.8%; Pred. No. 1.7e-216;	
	Matches 945; Conservative 0; Mismatches 0; Indels 96; Gaps 1;	
QY	1 atggaaaagtatgaaaaattagctaagactggagaagggtcttattgggttgattcaaa	60
Db		
QY	1 atggaaaagtatgaaaaattagctaagactggagaagggtcttattgggttgattcaaa	60
Db		
QY	61 tgcagaaaacacccctctggacaagtagtctgttaaaaaatttgggaatctgaagat	120
Db		
QY	121 gatcctgtgttaagaaaaatagcactaagagaatacgtatgttgaagcaattaaacat	180
Db		
QY	181 ccaaatcttgaacctcatcgagggtgttcaggagaaaaaggaaaatgcatttagtttt	240
Db		
QY	181 ccaaatcttgaacctcatcgagggtgttcaggagaaaaaggaaaatgcatttagtttt	240
Db		
QY	241 gaatactgtgatcatacacttttaaatgagctggaaagaaacccaaatggagtgcgat	300
Db		
QY	241 gaatactgtgatcatacacttttaaatgagctggaaagaaacccaaatggagtgcgat	300
Db		
QY	301 ggagtgatcaaaaagcgtattatggcacaacacttcaagctcttaatttctgcataacat	360
Db		
QY	301 ggagtgatcaaaaagcgtattatggcacaacacttcaagctcttaatttctgcataacat	360
Db		
QY	361 aactgtattcacagagatataaaaacctgaaaaattcttaataactaagcaaggaataatc	420
Db		
QY	361 aactgtattcacagagatataaaaacctgaaaaattcttaataactaagcaaggaataatc	420
Db		
QY	421 aagatttgtgacttcgggttgccacaaaattctgattccaggagatgcctacacgattat	480
Db		
QY	421 aagatttgtgacttcgggttgccacaaaattctgattccaggagatgcctacacgattat	480
Db		
QY	481 gtactacagagatgtaccgagctcctgaactcttctgtgggagatactcagtaggttct	540
Db		
QY	481 gtactacagagatgtaccgagctcctgaactcttctgtgggagatactcagtaggttct	540
Db		
QY	541 tcagtcgatataatgggctattgtgtgttttgcagagctcctgacagggccagccactg	600
Db		
QY	541 tcagtcgatataatgggctattgtgtgttttgcagagctcctgacagggccagccactg	600
Db		
QY	601 tggcctgaaaaatcagatgtggaccaactttatctgataatcagaaacact-----	650
Db		
QY	601 tggcctgaaaaatcagatgtggaccaactttatctgataatcagaaacacttagtagacg	660
Db		
QY	651 -----	650
Db	661 gggtttcgcatgttgaccaggctgtctcgaaactcttgagctcaagtgatccacctgcc	720
QY	651 -----	684
Db	721 gtgcctctcaaaagtgtggaattacaggaaaaataatcccaagacatcaatcaatctt	780
QY	685 aaaagtaacgggtttttccatggcatcagtagtgcagagcctcctgacagggccagccact	744
Db		
QY	781 aaaagtaacgggtttttccatggcatcagtagtgcagagcctcctgacagggccagccact	840
Db		
QY	745 gaggaaaaagtctcagatgttcactctgtggtcctgaacttcatgaagggtgtctgaag	804
Db		
QY	841 gaggaaaaagtctcagatgttcactctgtggtcctgaacttcatgaagggtgtctgaag	900
Db		

QY	805 atgaatccagatgacagattaacctgttccaaactctggagagctcctactttgattct	864
Db		
QY	901 atgaatccagatgacagattaacctgttccaaactctggagagctcctactttgattct	960
Db		
QY	865 ttccaagaggcccaaatataaagaaaagcacgtaataagaggaagaacagaaacgcca	924
Db		
QY	925 caggtacttcgctcaaaagt 945	
Db		
QY	1021 caggtacttcgctcaaaagt 1041	
Db		
RESULT	5	
AAD03813		
ID	AAD03813 standard; cDNA; 1068 BP.	
XX		
AC	AAD03813;	
XX		
DT	19-JUN-2001 (first entry)	
XX		
DE	Human kinase cDNA #2.	
XX		
KW	Human; kinase; gene therapy; bioreactor; mental disorder;	
KW	biological disorder; ss.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	1..1068
FT		/*tag= a
FT		/product= "Human kinase #2"
FT		/note= "the coding region does not include stop codon"
FT		/partial
XX		
PN	WO200123579-A1.	
XX		
PD	05-APR-2001.	
XX		
PF	27-SEP-2000; 2000WO-US26621.	
XX		
PR	28-SEP-1999; 99US-0156511.	
XX		
PA	(LEXI-) LEXICON GENETICS INC.	
XX		
PI	Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;	
XX		
DR	WPI; 2001-266166/27.	
DR	P-PSDB; AAE00491.	
XX		
PT	New isolated human kinase polynucleotide useful for generating	
PT	antibodies, as reagents in diagnostic assays and for screening for	
PT	compounds useful for treating mental, biological or medical diseases	
XX		
PS	Disclosure; Page 28; 38pp; English.	
XX		
CC	The present sequence is a cDNA encoding novel human protein	
CC	(NHP) known as human kinase. The human kinases share structural	
CC	similarity with animal kinases, more particularly serine or	
CC	threonine protein kinases. Human kinase cDNA is useful for the	
CC	detection of mutant human kinase for the diagnosis of disease,	
CC	and also as a therapeutic. It is useful for screening drugs	
CC	effective in the treatment of symptomatic or phenotypic	
CC	manifestations perturbing the normal function of NHP in the	
CC	body. The NHP nucleotide sequences are useful for generation of	
CC	antibodies, as reagents in diagnostic assays, for the	
CC	identification of other cellular gene products related to human	
CC	kinases, and as reagents in assays for screening compounds that	
CC	are useful for treating mental, biological or medical disorders.	
CC	NHP oligonucleotides are used as probes. The labelled NHP probes	
CC	are useful for screening human genomic library for identifying	
CC	polymorphisms and as primers in amplification assays to detect	
CC	mutations within the exons, introns and splice sites that can	

CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.
XX
SQ Sequence 1068 BP; 344 A; 197 C; 240 G; 287 T; 0 other;

Query Match 87.1%; Score 823; DB 22; Length 1068;
Best Local Similarity 90.6%; Pred. No. 3.5e-212;
Matches 929; Conservative 0; Mismatches 0; Indels 96; Gaps 1;

QY 1 atggaaaagtatgaaaaattagctaagactggagaagggtcttattgggttgattcaaa 60
Db |||||||
QY 61 tgcagaaacaaacctctggacaagttagtctgtttaaaaaatttgtggaatctgaagat 120
Db |||||||
QY 121 gatactgttgaagaaaatagcactaagagaaatacgtatgttgaagcaattaaaaacat 180
Db |||||||
QY 181 ccaaatotttgaacctcatcgaggtgttcaggagaaaaaggaaaatgcattagtttt 240
Db |||||||
QY 241 gaatactgtgatcatacacttttaaatgagctggaaagaaacccaaatggagttgctgat 300
Db |||||||
QY 301 ggagtgatcaaaagcgtattatggcaaacacttcaagctcttaatttctgtcatatacat 360
Db |||||||
QY 361 aactgtattcacagagatataaaacctgaaaataattcttaactaagcaaggaataatc 420
Db |||||||
QY 421 aagatttgtgacttcgggttgcacaaattctgattccaggagatgcctacacccgattat 480
Db |||||||
QY 481 gtagctacgagatggtaccgagctcctgaaattctgattccaggagatgcctacacccgattat 540
Db |||||||
QY 541 tcagtcgatatatgggctattggttgtgttttgcagagctcctgacaggtccctgagatggttct 600
Db |||||||
QY 601 tggcctggaaaaatcagatgtggaccaactttatctgataatcagaaacact----- 650
Db |||||||
QY 651 tggcctggaaaaatcagatgtggaccaactttatctgataatcagaaacacttagagacg 660
Db -----
QY 661 gggtttcgccatgttgaccaggctggtctcgaaactcttgacgtcaagtgatccacctgcc 720
QY 651 -----aggaaaaattaatcccaagacatcaatcaatcttt 684
Db |||||||
QY 721 gtagctctcaaaagtctgggaattacaggaaaaattaatcccaagacatcaatcaatcttt 780
QY 685 aaaagtaacgggttttccatggcatcagtatacctgagccagaagacatggaaaacttt 744
Db |||||||
QY 781 aaaagtaacgggttttccatggcatcagtatacctgagccagaagacatggaaaacttt 840
QY 745 gaggaagaagtctcagatgttcacacctgtggtctgaaactcatgaagggtgtctgaag 804
Db |||||||
QY 841 gaggaagaagtctcagatgttcacacctgtggtctgaaactcatgaagggtgtctgaag 900

QY 805 atgaatccagatgacagattaaacctgtgtcccaactcctggagagctcctacttattct 864
Db |||||||
QY 901 atgaatccagatgacagattaaacctgtgtcccaactcctggagagctcctacttattct 960
Db |||||||
QY 865 ttccaagagggcccaattaaaaagaaagcagcgtaatgaaggaagaaacagagacgcaa 924
Db |||||||
QY 961 ttccaagagggcccaattaaaaagaaagcagcgtaatgaaggaagaaacagagacgcaa 1020
QY 925 caggt 929
Db |||||
QY 1021 caggt 1025

RESULT 6
AAF44669
ID AAF44669 standard; cDNA; 911 BP.
XX
AC AAF44669;
XX
DT 27-MAR-2001 (first entry)
XX
DE Novel protein kinase cDNA, SEQ ID NO: 49.
XX
KW Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
KW immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic;
KW dermatological; antidiabetic; antinfertility; gene therapy; vaccine;
KW immune disorder; cardiovascular disease; neurodegenerative disease;
KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
KW inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.
XX
OS Homo sapiens.
XX
PN WO200073469-A2.
XX
PD 07-DEC-2000.
XX
PF 26-MAY-2000; 2000WO-US14842.
XX
PR 28-MAY-1999; 99US-0136503.
XX
PA (SUGE-) SUGEN INC.
XX
PI Plowman GD, Martinez R, Whyte D, Sudersanam S;
XX
DR WPI; 2001-032161/04.
DR P-PSDB; AAB65642.
XX
PT Nucleic acids encoding kinase polypeptides, useful for diagnosing and
PT treating immune-related diseases and disorders, cardiovascular disease,
PT neurodegenerative diseases and/or cancers -
XX
PS Example 4; Fig 2; 310pp; English.
XX
CC The present sequence encodes a novel protein kinase. The nucleic acids
CC and the protein kinases they encode may be used in the treatment and
CC diagnosis of diseases associated with inappropriate kinase expression
CC such as immune-related diseases and disorders, cardiovascular disease,
CC neurodegenerative diseases and/or cancers. The nucleic acids and
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays. The kinase polypeptides may be used as antigens in the production
CC of antibodies of kinase expression and activity. Anti-kinase antibodies
CC and kinase antagonists may also be used to down regulate kinase
CC expression and activity. Diseases related to kinase expression and
CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
CC disorders, complications of organ transplantation, myocardial infarction,
CC immune disorders, cardiomyopathies, strokes, renal failure,
CC oxidative-stress related disorders, chronic inflammatory bowel disease,
CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
CC reproductive disorders.
XX
SQ Sequence 911 BP; 293 A; 182 C; 178 G; 258 T; 0 other;

Query Match		68.0%;	Score 643;	DB 22;	Length 911;
Best Local Similarity		100.0%;	Pred. No. 1.1e-163;		
Matches 643;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	285	aaatggagttgctgagtgatgacaaaagcgtattatgagcaaacacttcaagctcttaa	344		
Db	1	aaatggagttgctgagtgatgacaaaagcgtattatgagcaaacacttcaagctcttaa	60		
QY	345	tttctgcataataaactgtattccacagagatataaaacctgaaaatatcttaataac	404		
Db	61	tttctgcataataaactgtattccacagagatataaaacctgaaaatatcttaataac	120		
QY	405	taagcaaggaaataatcaagatttgtgacttcgggtttgcacaaaattctgattccaggaga	464		
Db	121	taagcaaggaaataatcaagatttgtgacttcgggtttgcacaaaattctgattccaggaga	180		
QY	465	tgctacaccgattatgtagctacgagatggtaccgagctcctgaactcttctgtggaga	524		
Db	181	tgctacaccgattatgtagctacgagatggtaccgagctcctgaactcttctgtggaga	240		
QY	525	tactcagtatggttcttcagtcgatatatgggctattggtgtgtttttgcagagctcct	584		
Db	241	tactcagtatggttcttcagtcgatatatgggctattggtgtgtttttgcagagctcct	300		
QY	585	gacaggccagccactgtggcctggaaaatcagatgtggaccaactttatctgataatcag	644		
Db	301	gacaggccagccactgtggcctggaaaatcagatgtggaccaactttatctgataatcag	360		
QY	645	aacactaggaaaaattaatcccaagacatcaatcaatctttaaaaagtaacgggtttttcca	704		
Db	361	aacactaggaaaaattaatcccaagacatcaatcaatctttaaaaagtaacgggtttttcca	420		
QY	705	tggeatcagtatacctgagccagaagacatggaacctcttgaggaaaaagtctcagatgt	764		
Db	421	tggeatcagtatacctgagccagaagacatggaacctcttgaggaaaaagtctcagatgt	480		
QY	765	tcatcctgtggctctgaacttcagtgaaaggggtgtctgaagatgaatccagatgacagatt	824		
Db	481	tcatcctgtggctctgaacttcagtgaaaggggtgtctgaagatgaatccagatgacagatt	540		
QY	825	aacctgttcccaactcctggagagctcctactttgtattcttttcaagaggcccaaataa	884		
Db	541	aacctgttcccaactcctggagagctcctactttgtattcttttcaagaggcccaaataa	600		
QY	885	aagaaaagcacgtaataatgaaggagaagaacagagacgcgccaacag	927		
Db	601	aagaaaagcacgtaataatgaaggagaagaacagagacgcgccaacag	643		
RESULT 7					
AAF44670					
ID	AAF44670 standard; cDNA; 2615 BP.				
XX					
AC	AAF44670;				
XX					
DT	27-MAR-2001 (first entry)				
XX					
DE	Novel protein kinase cDNA, SEQ ID NO: 50.				
XX					
KW	Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;				
KW	immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic;				
KW	dermatological; antidiabetic; antiinfertility; gene therapy; vaccine;				
KW	immune disorder; cardiovascular disease; neurodegenerative disease;				
KW	cancer; autoimmune disorder; stroke; inflammatory bowel disease;				
KW	inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.				
XX					
OS	Mus musculus.				
XX					
PN	WO200073469-A2.				
XX					
PD	07-DEC-2000.				

X	26-MAY-2000; 2000WO-US14842.				
X	28-MAY-1999; 99US-0136503.				
X	(SUGE-) SUGEN INC.				
X	Plowman GD, Martinez R, Whyte D, Sudersanam S;				
I	WPI; 2001-032161/04.				
R	P-PSDB; AAB65643.				
R	Nucleic acids encoding kinase polypeptides, useful for diagnosing and				
X	treating immune-related diseases and disorders, cardiovascular disease,				
PT	neurodegenerative diseases and/or cancers -				
PT					
XX	Disclosure; Fig 2; 310pp; English.				
XX	The present sequence encodes a novel protein kinase. The nucleic acids				
CC	and the protein kinases they encode may be used in the treatment and				
CC	diagnosis of diseases associated with inappropriate kinase expression				
CC	such as immune-related diseases and disorders, cardiovascular disease,				
CC	neurodegenerative diseases and/or cancers. The nucleic acids and				
CC	complementary sequences may also be used as DNA probes in diagnostic				
CC	assays. The kinase polypeptides may be used as antigens in the production				
CC	of antibodies of kinase expression and activity. Anti-kinase antibodies				
CC	and kinase antagonists may also be used to down regulate kinase				
CC	expression and activity. Diseases related to kinase expression and				
CC	activity include rheumatoid arthritis, atherosclerosis, myocardial infarction,				
CC	disorders, complications of organ transplantation, strokes, renal failure,				
CC	immune disorders, cardiomyopathies, chronic inflammatory bowel disease,				
CC	oxidative-stress related disorders, multiple sclerosis, asthma,				
CC	chronic inflammatory pelvic disease, osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and				
CC	reproductive disorders.				
XX					
SQ	Sequence 2615 BP; 742 A; 580 C; 585 G; 708 T; 0 other;				
	Query Match 59.2%; Score 559.2; DB 22; Length 2615;				
	Best Local Similarity 81.8%; Pred. No. 6.3e-141;				
	Matches 658; Conservative 0; Mismatches 143; Indels 3; Gaps 1				
QY	135 gaaaatagcactaagagaaaatcgtatgttgaaagcaattaaacatccaaaatcttgtgaa	194			
Db	1 gaaaatagccctgcgggaaaatccgtatgctgtgaag---ttgaaacacccaacacctgtgaa	57			
QY	195 cctcatcgagggtgttcaggagaaaaaggaaaaatgcattttagtttttgaatactgtgatca	254			
Db	58 cctcatcgagggtgttcagaagaagaagaaaagatgcatactagttttttagtactgtgatca	117			
QY	255 tacacttttaaatgagctggaaagaaacccaaaatggagttgctgtagggagtgatcaaaaag	314			
Db	118 cacactgttaaacgagctggagagaaaccccaacggagtttctgtaggagtgattaaaag	177			
QY	315 cgtattatggcaaacacttcaagcttctaatttctgtcatatatacataactgtattcacag	374			
Db	178 tgtgctatggcaaaccttcaagcccttaacttctgtcacaagcacaaattgtattcatcg	237			
QY	375 agatataaaacctgaaaatatattctaactaagaagcaaggaataatcaagatttgtgactt	434			
Db	238 ggatgtaaaacctgaaaacatcctaataaccaagcaaggatgataaagatttgtgactt	297			
QY	435 cgggtttgcacaaaattctgattccaggagatgcctacaccgattatgtagctacagatg	494			
Db	298 tggatttgacgaattcttaattccaggagacgcctacacagactatgttgccaccagggtg	357			
QY	495 gtaccgagctcctgaactcttctgtgggagatactcagtatggttcttcagtcgatatatg	554			
Db	358 gtaccgagcccccgaaacttctcgtgggagacacgaagtacggttctctctgtagacgtgtg	417			
QY	555 ggctattggtgtgtttttgcagagctcctgcagagccagccactgtggcctggaaaatc	614			

Db 418 ggcgcgtcggctgtgttttgcagagctcctgcagcgggtcagccactctgcccgggaaaaatc 477

QY 615 agatgtggaccaactttatctgataatcagaaacactaggaataattatcccaagacatca 674

Db 478 cgacgtggaccagctttacctgatcatcaggacgttgggaagctgattccaagcacca 537

QY 675 atcaaatctttaaaagtaacgggttttccatggcatcagtcatacctgagccagaagacat 734

Db 538 gtctatctttaggagtaaccagttttccgcggcatcagcatcacctgaaccagaggacat 597

QY 735 ggaacctcttgaggaaaagtctcagatgttcatcctctgtgctctgaacttcataagg 794

Db 598 ggagactcttgaagaaaaattctcaaatgttcagcctgtggtctttaagtcttcataagg 657

QY 795 gtgtctgaagatgaatccagatgacagattaaacctgttccaaacctcctggagagctccta 854

Db 658 atgcctgaagatgaatcctgatgagagcgtgacctgtgccagcgtcgtggacagtgccta 717

QY 855 ctttgattcttttcaagagggcccaaatataaaagaaagcacgtaatgaaggagaacag 914

Db 718 ctttgagctctttcaagaggatcaaatgaaagaaagcccgacgtgaggggagaagcgg 777

QY 915 aagacgccaaacaggtacttcgcgt 938

Db 778 aaggcgcagcagaatcaactgct 801

RESULT 8

AAD03818

ID AAD03818 standard; cDNA; 1819 BP.

XX

AC AAD03818;

XX

DT 19-JUN-2001 (first entry)

XX

DE Human kinase cDNA #7.

XX

KW Human; kinase; gene therapy; bioreactor; mental disorder;

KW biological disorder; polymorphism; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT 5'UTR 1..137

FT /*tag= a

FT CDS 138..701

FT /*tag= b

FT /product= "Human kinase #1"

FT 3'UTR 702..1819

FT /*tag= c

FT misc_difference 1684

FT /*tag= d

FT /note= "polymorphism site; given as r in the sequence"

XX

PN WO200123579-A1.

XX

PD 05-APR-2001.

XX

PF 27-SEP-2000; 2000WO-US26621.

XX

PR 28-SEP-1999; 99US-0156511.

XX

PA (LEXI-) LEXICON GENETICS INC.

XX

PI Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;

XX

DR WPI; 2001-266166/27.

DR P-PSDB; AAE00490.

XX

PT New isolated human kinase polynucleotide useful for generating

PT antibodies, as reagents in diagnostic assays and for screening for

PT compounds useful for treating mental, biological or medical diseases

XX

PS Disclosure; Page 34-35; 38pp; English.

XX

CC The present sequence encodes a novel human protein (NHP) which

CC is a human kinase. A polymorphism was identified in the 3' UTR

CC of the present sequence. The human kinases share structural

CC similarity with animal kinases, more particularly serine or

CC threonine protein kinases. Human kinase cDNA is useful for the

CC detection of mutant human kinase for the diagnosis of disease,

CC and also as a therapeutic. It is useful for screening drugs

CC effective in the treatment of symptomatic or phenotypic

CC manifestations perturbing the normal function of NHP in the

CC body. The NHP nucleotide sequences are useful for generation of

CC antibodies, as reagents in diagnostic assays, for the

CC identification of other cellular gene products related to human

CC kinases, and as reagents in assays for screening compounds that

CC are useful for treating mental, biological or medical disorders.

CC NHP oligonucleotides are used as probes. The labelled NHP probes

CC are useful for screening human genomic library for identifying

CC polymorphisms and as primers in amplification assays to detect

CC mutations within the exons, introns and splice sites that can

CC be used in diagnostics and pharmacogenomics. Nucleotide construct

CC encoding NHP products are used to genetically engineer cells

CC in vivo that functions as bioreactors in the body delivering a

CC continuous supply of NHP to the body. Nucleotide constructs

CC encoding functional NHPs are used in gene therapy for the

CC modulation of NHP expression.

XX

SQ Sequence 1819 BP; 539 A; 368 C; 406 G; 505 T; 1 other;

Query Match 53.7%; Score 507; DB 22; Length 1819;

Best Local Similarity 69.8%; Pred. No. 6.5e-127;

Matches 929; Conservative 0; Mismatches 0; Indels 402; Gaps 2;

QY 1 atggaaaaagtatgaaaaattagctaagactggagaagggtcttatgggtgtgtattcaaa 60

Db 138 atggaaaaagtatgaaaaattagctaagactggagaagggtcttatgggtgtgtattcaaa 197

QY 61 tgcagaaacaaacacctctcggacaagttagctgtttaaaaaatttgtgaatctgaagat 120

Db 198 tgcagaaacaaacacctctcggacaagttagctgtttaaaaaatttgtgaatctgaagat 257

QY 121 gatcctgttgttaagaaaaatagcactaagagaaatacgtgttgaaagcaattaaacat 180

Db 258 gatcctgttgttaagaaaaatagcactaagagaaatacgtgttgaaagcaattaaacat 317

QY 181 ccaaatcttgtgaacctcatcgagggtgttcaggagaaaaaggaaaatgcatttagtttt 240

Db 318 ccaaatcttgtgaacctcatcgagggtgttcaggagaaaaaggaaaatgcatttagtttt 377

QY 241 gaatactgtgatacacacttttaaatgagctggaaagaaaccccaaatggagttgctgat 300

Db 378 gaatactgtgatacacacttttaaatgagctggaaagaaaccccaaatggagttgctgat 437

QY 301 ggagtgatcaaaagcgtattatggcaaacacttcaagctcttaatttctgtcatacat 360

Db 438 ggagtgatcaaaagcgtattatggcaaacacttcaagctcttaatttctgtcatacat 497

QY 361 aactgtattcacagagatataaaacctgaaaaatattctaataactaagcaaggaataatc 420

Db 498 aactgtattcacagagatataaaacctgaaaaatattctaataactaagcaaggaataatc 557

QY 421 aagatttgtgacttcgggtttgcacaaattctg----- 453

Db 558 aagatttgtgacttcgggtttgcacaaattctgagttggacttcattctctctgtgccc 617

QY 454 ----- 453

Db 618 tccttgattggccttaatagttgaccttctgaattcttttctgccaattcagagattttt 677

QY 454 ----- 453

Db 678 ctcttggttgatccattgtgacacagtgtttcaccatggggcccggtcatctcoga 737

Db 241 gaatactgtgatcatacacacttttaaatgagctggaagaaaccccaaatggagttgctgat 300
QY 301 ggagtgatcaaaagcgtattatggcaaacacttcaagctcttaatttctgtcatatacat 360
Db 301 ggagtgatcaaaagcgtattatggcaaacacttcaagctcttaatttctgtcatatacat 360
QY 361 aactgtattcacagagatataaaaacctgaaaatatttctaataactaagcaagggaataac 420
Db 361 aactgtattcacagagatataaaaacctgaaaatatttctaataactaagcaagggaataac 420
QY 421 aagatttggacttcgggtttgcacaaaattctgatt 456
Db 421 aagatttggacttcgggtttgcacaaaattctgatt 456

RESULT 10
AAD03815
ID AAD03815 standard; cDNA; 594 BP.
XX
AC AAD03815;
XX
DT 19-JUN-2001 (first entry)
XX
DE Human kinase cDNA #4.
XX
KW Human; kinase; gene therapy; bio reactor; mental disorder;
KW biological disorder; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..594
FT /*tag= a
FT /product= "Human kinase #4"
FT /note= "The coding region does not include stop codon"
FT /partial
XX
PN WO200123579-A1.
XX
XX
PD 05-APR-2001.
XX
XX
PF 27-SEP-2000; 2000WO-US26621.
XX
PR 28-SEP-1999; 99US-0156511.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
XX
DR WPI; 2001-266166/27.
XX
DR P-PSDB; AAE00493.
XX
PT New isolated human kinase polynucleotide useful for generating
PT antibodies, as reagents in diagnostic assays and for screening for
PT compounds useful for treating mental, biological or medical diseases -
XX
PS Disclosure; Page 31; 38pp; English.
XX
CC The present sequence is a cDNA encoding novel human protein
CC (NHP) known as human kinase. The human kinases share structural
CC similarity with animal kinases. The human kinases share structural
CC threonine protein kinases, more particularly serine or
CC detection of mutant human kinase for the diagnosis of disease,
CC and also as a therapeutic. It is useful for screening drugs
CC effective in the treatment of symptomatic or phenotypic
CC manifestations perturbing the normal function of NHP in the
CC body. The NHP nucleotide sequences are useful for generation of
CC antibodies, as reagents in diagnostic assays, for the
CC identification of other cellular gene products related to human
CC kinases, and as reagents in assays for screening compounds that
CC are useful for treating mental, biological or medical disorders.
CC NHP oligonucleotides are used as probes. The labelled NHP probes
CC are useful for screening human genomic library for identifying

CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bio reactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.
XX
SQ Sequence 594 BP; 203 A; 94 C; 115 G; 182 T; 0 other;

Query Match 48.1%; Score 454.4; DB 22; Length 594;
Best Local Similarity 99.8%; Pred. No. 6.4e-113;
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 atggaaaagtatgaaaaattagctaagactggagaagggtcttattggggtgtattcaaa 60
Db 1 atggaaaagtatgaaaaattagctaagactggagaagggtcttattggggtgtattcaaa 60
QY 61 tgcagaaaacaaaacctctggacaagttagctgtttaaaaaatttggaaatctgaagat 120
Db 61 tgcagaaaacaaaacctctggacaagttagctgtttaaaaaatttggaaatctgaagat 120
QY 121 gatcctgttgttaagaaaaatagcactaagagaaaatcgtatgttgaagcaattaaaacat 180
Db 121 gatcctgttgttaagaaaaatagcactaagagaaaatcgtatgttgaagcaattaaaacat 180
QY 181 ccaaatcttgtgaacctcctcaggtgttcaggagaaaaaggaatgcattagtttt 240
Db 181 ccaaatcttgtgaacctcctcaggtgttcaggagaaaaaggaatgcattagtttt 240
QY 241 gaatactgtgatcatacacacttttaaatgagctggaagaaaccccaaatggagttgctgat 300
Db 241 gaatactgtgatcatacacacttttaaatgagctggaagaaaccccaaatggagttgctgat 300
QY 301 ggagtgatcaaaagcgtattatggcaaacacttcaagctcttaatttctgtcatatacat 360
Db 301 ggagtgatcaaaagcgtattatggcaaacacttcaagctcttaatttctgtcatatacat 360
QY 361 aactgtattcacagagatataaaaacctgaaaatatttctaataactaagcaagggaataac 420
Db 361 aactgtattcacagagatataaaaacctgaaaatatttctaataactaagcaagggaataac 420
QY 421 aagatttggacttcgggtttgcacaaaattctgatt 456
Db 421 aagatttggacttcgggtttgcacaaaattctgatt 456

RESULT 11
AAA29746
ID AAA29746 standard; DNA; 1698 BP.
XX
AC AAA29746;
XX
DT 15-AUG-2000 (first entry)
XX
DE Rabbit KKIAMRE kinase genomic DNA SEQ ID NO:5.
XX
KW Rabbit; KKIAMRE kinase; learning-induced kinase; learning; memory;
KW cdc2-related kinase; brain; gene therapy; genetic disorder; detection;
KW identification; ds.
XX
OS Oryctolagus cuniculus.
XX
PN WO200020567-A2.
XX
PD 13-APR-2000.
XX
PF 01-OCT-1999; 99WO-US23010.
XX
PR 02-OCT-1998; 98US-0102906.
XX

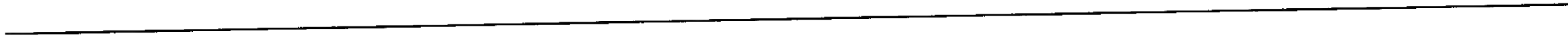
QY	181	ccaaatctgtgaacctcatcgagggtgttcaggagaaaaaggaaaaatgcatttagtttt	240
Db	178	gaaaacctgggtcaatctgattgaagtttttagacagaaaaaataattcatttgggtattt	237
QY	241	gaatactgtgatcacatacacttttaaatgagctggaagaaaccccaaatggagttgctgat	300
Db	238	gaatttattgaccacacagattagatgagttacaacattattgtcatggactagagagt	297
QY	301	ggagtgatcaaaaagcgtattattggcaaacacttcaagctctttaatttctgtcatatacat	360
Db	298	aagcgacttagaaaaataacctcttccagatccttcgagcaattgactatcttcacagtaat	357
QY	361	aactgtattcacagagatatataaaacctgaaaaatatttctaataactaagcaagggaataatc	420
Db	358	aatatcattcatcgagatatataaaacctgagaataatttttagtatccccagtcaggaattact	417
QY	421	aagatttgtgaacttcgggtttgcacaaaattct---gattccaggagatgcctacacogat	477
Db	418	aagctctgtgattttgtgtttgcagcaacactgacagctcctctgggacatttatacggac	477
QY	478	tatgtagctacgagatggtaccgagctcctgaacttcttgtgttttgcagagctcctgacaggccagcca	537
Db	478	tatgtggccacacgctggtatagagctcccgaaattagattataaagatacttcttatgga	537
QY	538	tcttcagtcgatatatgggctattgtgtgttttgcagagctcctgacaggccagcca	597
Db	538	aaacctgtggatatctgggctttgggctgtatgatcattgagatggccactggaaatccc	597
QY	598	ctgtggcctggaaaaatcagatgtggaccaacttttatctgataatcagaaaactaggaanaa	657
Db	598	tatcttctctagtagttctgattgttgattacttactocataaaaattgtttgaaagtgggcaat	657
QY	658	ttaatcccaagacatcaatcaatcttttaaaaagtaacggggtttttccatggcatcagtata	717
Db	658	ttgtcacctcacttgcagaatatcttttccaagagcccccatttttgcgtgggtagttctt	717
QY	718	cctgagccagaaagacatggaactcttgaggaaaaagttctcagatgttctcctgtggct	777
Db	718	cctcaagttcaacaccccccaaaaatgcaagaaaaaataatccaaagcttaattgattgttg	777
QY	778	ctgaacttcctgaaggggtgtctgaaagatgaatccagatgacagattaacctgttcccaa	837
Db	778	gcagatatagttcatgctgtgttttacaaaattgattcctctgctgacaggatatcatctagtat	837
QY	838	ctcctggagagctcctactttt	858
Db	838	cttttgcatcatgagtagtatttt	858
RESULT 14			
AAS06724			
ID	AAS06724 standard; cDNA; 1773 BP.		
XX			
AC	AAS06724;		
XX			
DT	12-SEP-2001 (first entry)		
XX			
DE	Polynucleotide sequence encoding human protein kinase #24.		
XX			
KW	Human; protein kinase; PTK; STK; cancer; cardiovascular disease;		
KW	metabolic disorder; immune related disease; neurological disorder;		
KW	neurodegenerative disorder; inflammatory disorder; infectious disease;		
KW	reproductive disorder; gene therapy; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200138503-A2.		
XX			
PD	31-MAY-2001.		
XX			
PF	22-NOV-2000; 2000WO-US32085.		
XX			
PR	24-NOV-1999; 99US-0167482.		

XX	(SUGE-) SUGEN INC.		
PA			
XX	Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;		
PI	Flanagan P, Clary D;		
XX	WPI; 2001-343950/36.		
DR	P-PSDB; AAU03524.		
XX	Nucleic acids encoding human kinase polypeptides, useful for preventing		
PT	diagnosing and/or treating e.g. cancer, immune, cardiovascular and		
PT	neuronal-associated diseases, and microbial infections -		
XX	Example 1; Figure 1; 433pp; English.		
PS			
XX	AAS06701-AAS06757 encode for novel human protein kinases #1-57. The		
CC	novel protein kinases have been identified as members of the tyrosine		
CC	or serine/threonine kinase (PTK and STK) families. The polynucleotides		
CC	encoding protein kinases and the polypeptides may be used in the		
CC	prevention, diagnosis and treatment of diseases associated with		
CC	inappropriate kinase expression. For example, they may be used to treat		
CC	cancers (especially cancers of haematopoietic origin), cardiovascular		
CC	disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),		
CC	immune related diseases (e.g. rheumatoid arthritis), neurological		
CC	disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.		
CC	Parkinson's disease), inflammatory disorders (e.g. asthma), infectious		
CC	disease (e.g. HIV) and reproductive disorders (e.g. infertility).		
CC	Additionally, polynucleotides encoding protein kinases may be		
CC	used for gene therapy and as DNA probes in diagnostic assays.		
CC	The protein kinase polypeptides may be used as antigens in the production		
CC	of antibodies against the protein kinases and in assays to identify		
CC	modulators of protein kinase expression and activity.		
XX	Sequence 1773 BP; 650 A; 300 C; 362 G; 461 T; 0 other;		
SQ			
Query Match 24.1%; Score 227.4; DB 22; Length 1773;			
Best Local Similarity 55.6%; Pred. No. 1.5e-51;			
Matches 479; Conservative 0; Mismatches 376; Indels 6; Gaps 2;			
QY	1	atggaaaagtatgaaaaaattagctaagacttggagagagggtctttatggggttgattcaaa	60
Db	1	atgggagatgtatgaaaccccttgaaaaagtgggagagggttacggaacagtcagtaa	60
QY	61	tgcagaaaaacaaacacctctggacaagtagtagctgttaaataatttgggaatctgaagat	120
Db	61	tgtaaacataagaataactgggcagatagtgccattaaagattttatgagagaccagaa	120
QY	121	gatcctgtgtttaagaaaaatagcactaagagaaatcgtatgttgaagcaatataaacat	180
Db	121	caatctgtc--aacaaaattcgcgatgagagaaataaagtcttctaaagcaatttcacac	177
QY	181	ccaaatcttgtgaacctcatcgaggtgttcaggagaaaaaggaaatgcatttagtttt	240
Db	178	gaaaaccttgtcaatctgattgaagtttttagacagaaaaagaaaattcatttgggtattt	237
QY	241	gaatactgtgatcatcacacttttaaatgagctggaaaaagaaacccaatggagttgctgat	300
Db	238	gaatttattgaccacacagattattagatgagttacaacattattgtcatggactagagagt	297
QY	301	ggagtgatcaaaaagcgtattattggcaaacacttcaagctcttaatttctgcataatacat	360
Db	298	aagcgacttagaaaaataacctcttccagatccttcgagcaattgactatcttcacagtaat	357
QY	361	aactgtattcacagagatatataaaccttgaaaaataatcttaataactaagcaagggaataatc	420
Db	358	aatatcattcatcgagatatataaacctcgagaataatttagtatcccagtcaggaattact	417
QY	421	aagatttgtgacttcgggtttgcacaaaattct---gattccaggagatgcctacacgat	477
Db	418	aagctctgtgattttgtgtttgcacgaacactagcagctcctctctctgtacatttacggac	477
QY	478	tatgtagctacgagatggtaccgagctcctgaacttcttgtgttttgcagagctcctgacaggccagcca	537

us-09-671-050-11.rng

· · · Fri May 3 11:02:07 2002

Search completed: May 2, 2002, 17:13:19
Job time: 5456 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 17:06:37 ; Search time 2235.62 Seconds
(without alignments)
4542.257 Million cell updates/sec

Title: US-09-671-050-11
Perfect score: 945
Sequence: 1 atggaaaagtatgaaaatt.....aggtacttcgctcaaaagt 945

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_htc:*
10: gb_est1:*
11: gb_est2:*
12: gb_htc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	401.8	42.5	1691	12 AK016781	AK016781 Mus muscu
2	333.6	35.3	500	10 AI385966	AI385966 ml34h09.y
3	316.8	33.5	938	10 AA061797	AA061797 ml34h09.f
4	258.6	27.4	731	10 AW106692	AW106692 um32d03.y
5	255	27.0	718	10 AA286088	AA286088 vc33b05.f
6	209.8	22.2	579	10 AW233105	AW233105 fj28b10.y
7	200.4	21.2	295	10 AA626859	AA626859 zu89f10.s
8	192.4	20.4	1535	12 BC010966	BC010966 Homo sapi
9	186.6	19.7	727	10 AI508835	AI508835 vc33b05.y
10	174.8	18.5	582	10 AV672492	AV672492 AV672492
11	174.4	18.5	665	13 AQ953706	AQ953706 Sheared D
12	170.6	18.1	399	10 BE700387	BE700387 QV3-NN009

13	165.6	17.5	401	10 BE700389	BE700389 QV3-NN009
14	158.4	16.8	495	10 AI385990	AI385990 ml50h06.y
15	155.4	16.4	384	10 BE223187	BE223187 kp73g08.y
16	154.6	16.4	744	11 BF787339	BF787339 602113731
17	154	16.3	879	11 BG281841	BG281841 602403067
18	154	16.3	926	10 AL524036	AL524036 AL524036
19	150.4	15.9	779	11 BG760621	BG760621 602717373
20	146	15.4	1112	10 AT002436	AT002436 AT002436
21	145.4	15.4	628	10 AW611190	AW611190 un78e06.y
22	142.8	15.1	808	10 AI746896	AI746896 ul08g10.y
23	141.4	15.0	957	11 BI410031	BI410031 602962512
24	141.2	14.9	795	10 BE275968	BE275968 601120739
25	136.2	14.4	754	11 BG757858	BG757858 602711380
26	135.2	14.3	699	10 AI584951	AI584951 AL584951
27	135.2	14.3	752	10 BE746563	BE746563 601580135
28	135.2	14.3	911	11 BG397188	BG397188 602434149
29	134.8	14.3	1131	10 BE563495	BE563495 601334976
30	134.2	14.2	764	11 BG162489	BG162489 df15e08.y
31	134	14.2	202	10 BE135229	BE135229 ug28e10.y
32	134	14.2	696	11 BF479244	BF479244 L48-2906T
33	133	14.1	680	10 AL585015	AL585015 AL585015
34	132.8	14.1	619	11 BE844514	BE844514 EST259 Ap
35	131.8	13.9	1000	11 BF316057	BF316057 601896036
36	131.4	13.9	831	11 BG825471	BG825471 602749632
37	131.2	13.9	884	11 BG440417	BG440417 GA_Ea000
38	131	13.9	297	10 AI965145	AI965145 fc87e07.y
39	131	13.9	613	10 AL504123	AL504123 AL504123
40	131	13.9	778	10 BE413170	BE413170 MCG014.CO
41	131	13.9	963	11 BG420755	BG420755 602448546
42	130.6	13.8	732	10 AW701949	AW701949 ug93b12.y
43	130.2	13.8	793	11 BG700342	BG700342 602680030
44	130.2	13.8	796	11 BG915878	BG915878 602815329
45	130	13.8	713	10 AI246661	AI246661 qi27c01.x

ALIGNMENTS

RESULT 1

AK016781
LOCUS AK016781 1691 bp mRNA HTC 05-JUL-2001
DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4933411017, full insert sequence.
ACCESSION AK016781
VERSION AK016781.1 GI:12855701
KEYWORDS CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male testis cDNA to mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library clone:4933411017.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1691)
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Methods in enzymology. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2 (bases 1 to 1691)
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome research. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3 (bases 1 to 1691)
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,

,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
Waterston,R. and Wilson,R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This read is a RESEQUENCE of a previously sequenced mouse clone
correct orientation)
This read has been verified (found to hit its original self in the
Seq primer: -40RP from Gibco
High quality sequence stop: 339.
Location/Qualifiers

FEATURES
source
1. .500
/organism="Mus musculus"
/strain="Inbred CD-1"
/db_xref="taxon:10090"
/clone="IMAGE:513953"
/clone_lib="Stratagene mouse testis (#937308)"
/sex="males"
/tissue_type="testis"
/dev_stage="10-12 week old"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: testis; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. Average insert size: 1.0 kb; Uni-ZAP XR Vector;
-5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor
sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"

BASE COUNT 145 a 114 c 121 g 120 t
ORIGIN
Query Match 35.3%; Score 333.6; DB 10; Length 500;
Best Local Similarity 82.3%; Pred. No. 1.5e-66;
Matches 408; Conservative 0; Mismatches 84; Indels 4; Gaps 2;

QY 135 gaaatagcactaagagaaaatacgtatgttgaagcaattaaacatccaaatcttgtgaa 194
||||| || | ||||| ||||| ||||| || ||||| ||||| || |||||
Db 2 GAAATAGCCCTGCGGAAATCCGTATGCTGAAG--TTGAAACACCCAAACCTCGTGAA 58
QY 195 cctcatcgaggtgttcaggagaaaaagaaatgcatttagtttgaatacttgtatca 254
||||| ||||| ||||| || || ||||| ||||| ||||| ||||| |||||
Db 59 CCTCATCGAGGTGTTCAGAAGAAGAGAAAGATGCATCTAGTGTTTTGAGTACTGTGATCA 118
QY 255 tacacttttaaatgagctggaaagaaacccaaatggagttgctgagtgcgtgatcaaaag 314
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 119 CACACTGTTAACGAGCTGGAGAGAAACCCAAACGGAGTTCTGTATGGAGTGATTAAAG 178
QY 315 cgtattatggcaaacacttcaagctctttaattctgtcatatatacataactgtattcacag 374
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 179 TGTGCTATGCAAAACCCCTCAAGCCCTTAACCTCTGTCAAGGCACAAATTGTATTTCATCG 238
QY 375 agatataaaacctgaaaaattcttaataactaagcaaggaataatcaagatttgtgactt 434
|| ||||| ||||| || ||||| ||||| ||||| || ||||| ||||| |||||
Db 239 GGATGTAAACCTGAAAAACATCCTAATAACCAAGCAAGGGATGATAAAGATTGTGACTT 298
QY 435 cgggtttgcacaaattctgattccaggagatgacctacacccgattatgtagctacgagatg 494
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 299 TGGATTTCACGAATTCATAATCCAGGAGACGCCCTACACAGACTATGTTGCCACCAGGTG 358
QY 495 gtaccgagctcctgaactctttgtgggagatactcagtatggttcttcagtcgatataatg 554
||||| || ||||| ||||| || ||||| ||||| ||||| ||||| |||||
Db 359 GTACCGAGCCCTCGAACTTCTCGTGGGAGACACGAAGTACGGTTCCTCTGTAGACGCTGTG 418
QY 555 ggctattggttgtgtttttgcagagctcctgcagagccagccactgtggcctggaaaaatc 614
||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 419 GGCCGTCGGCTGTGTTTATGCAGAGAGCTTCTGACGTGTGAGCCACTCTGGCCCGG-AAATA 477
QY 615 agatgtgagaccaactt 630
|| ||||| ||||| |||||
Db 478 CGACGTGACCACTT 493

RESULT 3
AA061797 938 bp mRNA EST 03-FEB-1997
LOCUS ml34h09.r1 Stratagene mouse testis (#937308) Mus musculus cDNA
DEFINITION clone IMAGE:513953 5', similar to gb:X66358 SERINE/THREONINE-PROTEIN
KINASE KKIALRE (HUMAN);, mRNA sequence.

ACCESSION AA061797
VERSION AA061797.1 GI:15555606
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 938)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.

TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:307801
Seq primer: -28ml3 rev1 ET from Amersham
High quality sequence stop: 442.

FEATURES
source
1. .938
/organism="Mus musculus"
/strain="Inbred CD-1"
/db_xref="taxon:10090"
/clone="IMAGE:513953"
/clone_lib="Stratagene mouse testis (#937308)"
/sex="males"
/tissue_type="testis"
/dev_stage="10-12 week old"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: testis; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. Average insert size: 1.0 kb; Uni-ZAP XR Vector;
-5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor
sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"

BASE COUNT 254 a 215 c 236 g 233 t
ORIGIN

Query Match 33.5%; Score 316.8; DB 10; Length 938;
Best Local Similarity 82.9%; Pred. No. 1.1e-62;
Matches 398; Conservative 0; Mismatches 77; Indels 5; Gaps 3;

QY 135 gaaatagcactaagagaaaatacgtatgttgaagcaattaaacatccaaatcttgtgaa 194
||||| || | ||||| ||||| ||||| || ||||| ||||| || |||||
Db 1 GAAATAGCCCTGCGGGAATCCGTATGCTGAAG--TTGAAACACCCAAACCTCGTGAA 57
QY 195 cctcatcgaggtgttcaggagaaaaagaaaaatgcatttagtttgaatacttgtatca 254
||||| ||||| ||||| || ||||| ||||| ||||| ||||| ||||| |||||
Db 58 CCTCATCGAGGTGTTCAGAGAGAAAGAGATGCATCTAGTGTTTTGAGTACTGTGATCA 117
QY 255 tacacttttaaatgagctggaaagaaacccaaatggagttgctgagtgcgtgatcaaaag 314

[illegible]

RESULT	4				
AW106692					
LOCUS		731 bp	mRNA	EST	20-OCT-1999
DEFINITION		um32d03.y1 Sugano mouse kidney mkia Mus musculus cDNA clone			
		IMAGE:2236229 5' similar to gb:X66358 SERINE/THREONINE-PROTEIN			
		KINASE KIALRE (HUMAN);, mRNA sequence.			
ACCESSION		AW106692			
VERSION		AW106692.1	GI:6077492		
KEYWORDS		EST.			
SOURCE		house mouse.			
ORGANISM		Mus musculus			
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;			
REFERENCE		1	(bases 1 to 731)		
AUTHORS		Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.			
TITLE		The WashU-NCI Mouse EST Project 1999			
JOURNAL		Unpublished (1999)			
COMMENT		Contact: Marra M/WashU-NCI Mouse EST Project 1999 Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watson.wustl.edu This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGI:1006441			

be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTAAAGCTGG and 3' end primer CGACCTGGAGCTCGAGACA."

BASE COUNT	202 a	179 c	191 g	157 t	2 others
ORIGIN	Primer CGACCCTGACGCTCCGAGCACCA.				
Query Match	27.4%; Score 258.6; DB 10; Length 731;				
Best Local Similarity	66.9%; Pred. No. 2.5e-49;				
Matches 382; Conservative	0; Mismatches 186; Indels 3; Gaps 1;				
QY	5	aaaagtatgaaaaaattagctaaagctggagaagggtcttatgggttggttatccaaatgca	64		
Db	156	AAATATGAAAAAANAATGGAAAGATTGGAGAAGGCTCCTATGCGGTACTGTTCAAGTGCA	215		
QY	65	gaacaaaaacctctggacaaagtagtagctgtgttaaaaaattgtggaatctgaagatgac	124		
Db	216	GAAACAGGGACACGGGTGAGATCTGGCCATCAAGAGGTTCTGGAAACCCGAAGATGACC	275		
QY	125	ctgtgtttaaaaaatagcactaagagaaatacgtatgttgaagcaattaaaaacatccaa	184		
Db	276	CTGTGATAAGAAAAATCGCCCTTCGAGAAATCGCATGCTCAAGCAACTCAAGCACCCCA	335		
QY	185	atctgtgaacctcaatcgaggtgttcaggagaaaaagaaaaatgcatttagttttgaat	244		
Db	336	ACCTGGTCAACCTCCTGGAAAGTCTCCGGAGGAAGCGGAGGCTTCACCTGGTGTTCGAGT	395		
QY	245	actgtgatcatcacacttttaaatgagctggaagaaacccaaatggagttgctgatggag	304		
Db	396	ACTCGGACCACACGGTGTCTCACGAGCTGGATCGGTATCAGAGGGGGGTACCGAGGCCTC	455		
QY	305	tgatcaaaagcgtattatggcaaacacttcaagcttcaagcttcaattctgtcatatacaaat	364		
Db	456	TCGTGAAGAAACATAAATTGGCAGACACTGCAGGCTGTAAATTTCTGCCATAAAACATAACT	515		
QY	365	gtattcacagagatataaaaacctgaaaatatattctaataactaaagcaaggataatacaaga	424		
Db	516	GCATACACAGAGACGTGAAGCCGGAAAAATATTCATATCACCRAACACAGTCACCCATTAAAGC	575		
QY	425	tttgtgactcgggttttgacaaaattctctgatt----ccaggagatgcctacaccgattatg	481		
Db	576	TCTGTGACTTTGGGGTCCGACACGGCTCCTCACTGGACCCANGTGACTACTACACAGACTACN	635		
QY	482	tagctacagagatggtaccgagctcctgaacctcttctgtgggagatactcagtatggttctt	541		
Db	636	GTGCCCCACCCGTGGTACTGCTCACCCGAGCTGCTAGTTGGACACACGCAGTATTGGTCCC	695		
QY	542	cagtcgatatatgggctattggttgtgttt	572		
Db	696	CTGTAGATGTCTGGGCAATTGGCTGTGTGTT	726		

RESULT	5				
AA286088					
LOCUS					
DEFINITION					
AA286088		718 bp	mRNA	EST	09-APR-1997
vc33b05.r1		Barstead	MPLRBI	Mus musculus	cDNA clone IMAGE:76337 5'
similar to gb:X66358		SERINE/THREONINE-PROTEIN KINASE	KKIALRE	(HUMAN	
);, mRNA sequence.					
AA286088					
AA286088.1		GI:1932198			
EST.					
house mouse.					
Mus musculus					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
1 (bases 1 to 718)					
Marra, R., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,					
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,					
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,					
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and					
ORGANISM					
REFERENCE					
AUTHORS					

QY 1 atggaaaagtatatgaaaaaattagctaagactgagagagggtcttattggttgattcaaa 60
||||| ||||| || | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 215 ATGGAGAACTATGAGAAGATCAGTAAGATTGGCGAAGGTCGTACGGCGTTGTGTCAAG 274

QY 61 tgcagaaaaaaacacctgtgacaaagtagtagtctgttataaaatttttggaatctgaagat 120
||||| || ||| | ||||| || ||| || ||||| ||||| ||||| ||||| ||||| |||||
Db 275 TGCAGGAATAAAGACACCGGACAGATTGTCGCCATCAAGAAGTTTGTGGAGTCTGAGGAT 334

QY 121 gatcctgtgtttaagaaaaatagcactaaagagaaatacgtatgttgaagcaattaaaaacat 180
||||| | ||||| ||||| || ||||| ||||| || ||||| ||||| ||||| ||||| |||||
Db 335 GATCCCATCATTAAGAAATAGCGCTCAGAGAAATCCGCATGCTGAAGCAACTAAACAT 394

QY 181 ccaaatctgtgaacctcatcgagggtgttcaggagaaaaaggaatgcattagttttt 240
||||| ||||| || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 395 CCAAACTGGTGAATCTGTATGAGGTTTCAGAGAAAGAGAAACTTCACCTGGTGT 454

QY 241 gaatactgtgatcacacttttaaatgagctggaagaaacccaaatggagttgctgat 300
||| ||||| || || | ||||| ||||| || ||||| ||||| || ||||| || ||||| ||
Db 455 GAGTGCTGAGACCACACTGTGCTGAATGAGCTGGACAGATACCCACGAGGTGTCCAGAG 514

QY 301 ggagtgatcaaaagcgtattatggcaaacacttcaagctottaattcttcatatatacat 360
||| ||||| || ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 515 CATATGGTTAAAGCATCATCTGGGCAACACTTTAGGCTGTGAACCTCTGCCACAAACA 574

QY 361 aactg 365
|| |||
Db 575 AATTG 579

RESULT 7
AA626859/c
LOCUS
DEFINITION
AA626859
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
AA626859 295 bp mRNA EST 15-OCT-1997
zu89f10.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:745195
3' similar to SW:KKIA_HUMAN Q00532 SERINE/THREONINE-PROTEIN KINASE
KKIALRE ;, mRNA sequence.
AA626859
AA626859.1 GI:2539246
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 295)
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
White,Y., Wylie,T., Waterston,R. and Wilson,R.
WashU-NCI human EST Project
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40ml3 fwd. Et from Amersham
High quality sequence stop: 292.
Location/Qualifiers
1. .295
/organism="Homo sapiens"
/db_xref="GDB:5933055"
/db_xref="taxon:9606"
/clone="IMAGE:745195"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCGGCCCAATTTTTTTTTTTT 3'].

Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo. "

95 a 69 c 55 g 76 t

Query Match 21.2%; Score 200.4; DB 10; Length 295;
Best Local Similarity 95.0%; Pred. No. 5.8e-36;
Matches 207; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 438 gtttgacaaaattctgtattccaggagatgcctacacgattatgttagctacgagatgta 497
||||| || ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 280 GATTGATGCATTTGCAGTTCACAGGAGATGCCTACACCGATTATGTAGCTACGAGATGGTA 221

QY 498 ccgagctcctgaacttctgtgtggagatactcagtcagtgtgttcttcagtcgatatatgggc 557
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 220 CCGAGCTCCTCAACTTCTTGTGGGAGATACTCAGTATGTTCTTCAGTCGATATATGGGC 161

QY 558 tattggttggttttttgagagctctgcacaggccacccactgtggcctggaaaaatcaga 617
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 160 TATTGGTTGTGTTTTTTCAGAGACTCTCGACAGGCCAGCCACTGTGGCCTGGAAAATCAGA 101

QY 618 tgtggaccaacttttatctgataatacagaacactaggaa 655
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 100 TGTGGACCAACTTTATCTGATAATCAGAACACTAGGTA 63

RESULT 8
BC010966
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK
COMMENT
FEATURES
source
BC010966 1535 bp mRNA HTC 25-JUL-2001
Homo sapiens, Similar to serine/threonine kinase 9, clone
IMAGE:4294753, mRNA.
BC010966
BC010966.1 GI:15012116
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 1535)
Strausberg,R.
Direct Submission
Submitted (23-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NTH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: CLONTECH
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbiology.org
contact: amadan@systemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
Greene, Mark Kettelman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 19 Row: 0 Column: 9
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 3559924
This clone has the following problem: incomplete processing.
Location/Qualifiers
1. .1535
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4294753"
/tissue_type="Prostate"


```
QY 241 gaatactgtgatcatacacactttttaaatgagctggaaagaaccccaaatggagttgctgat 300 *
Db 350 GAGTACTGGGACCACACCGGTGCTTCACAGCNTGGATCGGTATCAAGGGGGGTACCAGAG 409
QY 301 ggagtgatcaaaaagcgtattattatggcaaacacttcaagctctttaatttctgtcatatacat 360
Db 410 CTNCTCGTGAAGA-CATNNACTTGCAGACACTGCAGGCTGTNAATTCTTGGCATTAAACA 468
QY 361 aactgtattcacagagatatataaacctgaaatatcttaataactaagcaaggaaataatc 420
Db 469 TACTGCATACCAAGGACGCTGAAGCCGGGAAATATTCTCATCACAAAACAGTCNAGCATT 528
QY 421 aagatttgtgacttcgggtttgcac 445
Db 529 AAGCCCTGTATNNTTTGGGTTCGAC 553

RESULT 10
AV672492
LOCUS AV672492 582 bp mRNA EST 05-OCT-2000
DEFINITION AV672492 Nori Satoh unpublished cDNA library Clona intestinalis
cDNA clone citb5a12 5', mRNA sequence.
ACCESSION AV672492
VERSION AV672492.1 GI:10110491
KEYWORDS EST.
SOURCE Ciona intestinalis.
ORGANISM Ciona intestinalis.
REFERENCE 1 (bases 1 to 582)
AUTHORS Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
TITLE Phlebobranchia; Clonidae; Ciona.
JOURNAL Satoh,N., Satou,Y., Kohara,Y. and Shin-i,T.
COMMENT Expressed genes in Ciona intestinalis
Unpublished (2000)
Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satohe@ascidian.zool.kyoto-u.ac.jp.

FEATURES
source
1..582
/organism="Ciona intestinalis"
/db_xref="taxon:7719"
/clone="citb5a12"
/tissue_type="whole animal"
/dev_stage="tailbud"

BASE COUNT 178 a 100 c 143 g 159 t 2 others
ORIGIN

Query Match
Best Local Similarity 18.5%; Score 174.8; DB 10; Length 582;
Matches 323; Conservative 60.3%; Pred. No. 4.3e-30;

QY 1 atggaaaagtatgaaaaaattagctaagactggagaagggtctttatgggggttgattcaaa 60
Db 47 ATGGAGAAATATGACAACCTGGGCATGGTGGGTGAAGGCTCCTATGGGATGGTTATGAAA 106
QY 61 tgcagaacaaaacacctctggcaaatagtagctgttaaaaaattgtggaatctgaagat 120
Db 107 TGTAACATAGGAAAAGTGGCAAAATTGTGCCATTAGAAGTTCTGGAAAGTGAAGAT 166
QY 121 gatcctgtgttaagaaaaatagc-actaagagaaaatacgtatgttggaagcaattaaaaa 179
Db 167 GACAAGATGGTGAAAAAGATGCGATGCGTGGAAGTTCGAAGTTTGAGAAAACCTCCATCA 226
QY 180 tccaatcttgtgaacctcatcgaggtgttcaggagaaaaaagaaaatgcatttagtttt 239
Db 227 CGAAACCTCGTCAACTTAATTGAAGTTTCCGTCGACGTACCGTCTTTATCTCGTCTT 286
QY 240 tgaatactgtgatcatacacactttttaaatgagctggaaagaaccccaaatggagttgctga 299
```

```
Db 287 TGAATTTGTGGACCACACGGTGTCTGGATGACCTTGANAAATATCCTTAACGGATTAAACGA 346
QY 300 tggagtgatcaaaaagcgtattatggcaaacacacttcaagctctttaatttctgtcatataca 359
Db 347 GATGACAGTTCGAAAAAATTCTGTGGCAAGTTCTTAGAGGAGTAGAATTCTGTACACGCCA 406
QY 360 taactgtattcacagagatatataaacctgaaatatattctaataactaagcaaggaaataat 419
Db 407 TAATATCATACACAGGGATATCAAACTGAAATAATTTTAAACTCAGATCGGGGGTTCAT 466
QY 420 caagatttgtgacttcgggtttgcac---aaattctgattccaggagatgcctacaccca 476
Db 467 CAAACTATGTGACTTTGGTTTGTCTCGGACATTAGCTGCTCCTGGTGAGGTTTATACGGA 526
QY 477 ttatgtagctacgagatggtaccgagctcctgaactcttctgtggagagatactcagt 532
Db 527 TTATGTGGCCACTAGGTGGTATCGTGTGCGCCGAGCTGTGGTCGGAGACACCAAAAT 582

RESULT 11
AQ953706
LOCUS AQ953706 665 bp DNA GSS 27-JAN-2000
DEFINITION Sheared DNA-53L19.TF Sheared DNA Trypanosoma brucei genomic clone
cDNA-53L19, DNA sequence.
ACCESSION AQ953706
VERSION AQ953706.1 GI:6776971
KEYWORDS GSS.
SOURCE Trypanosoma brucei.
ORGANISM Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
Trypanosoma.
REFERENCE 1 (bases 1 to 665)
AUTHORS El-Sayed,N., Zhao,S., Zhao,H., Gill,S., Suh,E., Malek,J., Fujii,C.,
Gerrard,C., Leech,V., de Jong,P., Ullu,E., Melville,S., Donelson,J.,
Fraser,C. and Adams,M.
TITLE Determination of clone end sequences from Trypanosoma brucei GUTat
10.1 sheared DNA library
JOURNAL Unpublished (1999)
COMMENT other_GSSs: Sheared DNA-53L19.TR
Contact: Najib M. El-Sayed
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: nelsayed@tigr.org
Clones are derived from the Trypanosoma brucei GUTat 10.1 sheared
DNA library constructed at TIGR. Clones will be available for
distribution through ATCC. Sheared DNA end sequences search page:
http://www.tigr.org/tldb/mdb/tbdb/.
Seq primer: M13-Forward
Class: shotgun.

FEATURES
Location/Qualifiers
1..665
/organism="Trypanosoma brucei"
/strain="TREU927/4 GUTat 10.1"
/db_xref="taxon:5691"
/clone="Sheared DNA-53L19"
/clone_lib="Sheared DNA"
/note="Vector: pUC18; Site_1: SmaI; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from a cloned population of
Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically
sheared to give a tight size distribution (approx 2 kb).
The v + i method used for the library construction is
described in detail in Smith, H.O. and Venter, J.C.
(Making small insert libraries for whole genome shotgun
sequencing projects. In Genome Sequencing: A Practical
Approach, eds. M. Vaudin and B. Barell, Oxford University
Press, 1999)."
```

BASE COUNT 183 a 122 c 190 g 170 t
ORIGIN


```
QY 309 caaaagcgtattatggcaaacacttcaagctcttctaattctgtcatatatacataaactgtat 368
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 198 GAGAAAGTACCTCTTCCAGATCCTTCGAGCCATTTCAGTACCTGCATAATAATAATATATAT 257
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 369 tcacagagatataaaaacctgaaataattcttaataactaagcaaggaataataocaagatttg 428
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 258 CCATCGAGATATAAAGCCTGAGAAATATTTTAGTCTCCAGTCAGGAATTACGAAGCTGTG 317
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 429 tgacttcgggttgcacaaaattct---gattccaggagatgcctacacccgattatgtagc 485
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 318 TGATTTGGGTTTTCGAGGACACTAGCAGCTCCTCGAGACGTTTACACAGACTACGTGGC 377
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 486 tacgagatggtaccgagctcctgaaacttcttgtgggagatactcagtatggttcttcagt 545
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 378 CACACGCTGGTACAGAGCTCCAGAGCTGGTGTGAAAGACACCTNCTACGGANAGCCAGT 437
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 546 cgatatatgggctattggttgt 569
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 438 GGACATONTGGCTTTNGGCTGTAT 461
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 15
BE223187 BE223187 384 bp mRNA EST 09-MAY-2001
LOCUS kp73g08.y1 TBN95TM-SSFH Strongyloides stercoralis cDNA 5' similar
DEFINITION to WP:Y42A5A.4 CE20258 EUKARYOTIC PROTEIN KINASE DOMAIN ;, mRNA
sequence.
ACCESSION BE223187
VERSION BE223187.1 GI:8927762
KEYWORDS EST.
SOURCE Strongyloides stercoralis.
ORGANISM Strongyloides stercoralis
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Panagrolaimoidea; Strongyloidea; Strongyloides.
REFERENCE 1 (bases 1 to 384)
AUTHORS McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T.,
, Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y.,
, Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagareishvili,R.,
, Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe
,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S.,
, Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and
Wilson,R.
TITLE The Washington Univ. Nematode EST Project, 1999
JOURNAL Unpublished (1999)
COMMENT Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
The library was constructed by Dr. Thomas Nutman and colleagues of
NIAID, NIH (tnutman@nih.gov). DNA Sequencing by: Washington
University Genome Sequencing Center St. Louis.
Seq primer: -40RP from Gibco
High quality sequence stop: 340.
FEATURES
source Location/Qualifiers
1..384
/organism="Strongyloides stercoralis"
/strain="Filariform larvae obtained from humans"
/db_xref="taxon:6248"
/clone_lib="TBN95TM-SSFH"
/lab_host="XL-1 Blue MRF" (Stratagene)"
/note="Vector: Lambda Uni-ZAP XR (Stratagene); Site_1:
EcoRI; Site_2: XhoI; mRNA was purified from 4 x 10E5
filariform larvae which had been isolated from infected
humans. cDNA was constructed and, using adaptors, was
cloned unidirectionally into the vector from the EcoRI
site to the XhoI site. The library has an unamplified
titer of 1.5 x 10E6 pfu/ml and an amplified, undiluted
titer of 7 x 10E9 pfu/ml. The average insert size of the
unamplified library is 975 bp (range, 500-1500)."
```

BASE COUNT 120 a 37 c 81 g 144 t 2 others

```
ORIGIN
Query Match 16.4%; Score 155.4; DB 10; Length 384;
Best Local Similarity 62.7%; Pred. No. 1.2e-25;
Matches 240; Conservative 0; Mismatches 143; Indels 0; Gaps 0;
QY 287 atggagttgctgatggatgacaaaagggtattatggcaaacacttcaagctcttaatt 346
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2 AAGGATGTCCTGATGTTTGTATTAATAAACTATTTATCAATTTATGTTTAGCTATAAGAT 61
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 347 tctgtcatatacataaactgtattcacagagatatataaacctgaaaaatattcttaataacta 406
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 62 ATTGTCATAAACATAAATTGTATGCATAGAGATGTTAAACCAGAAAAATATTTTATTACAA 121
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 407 agcaaggaataatacaagatttgcacttcgggtttgcacaaaattcttgatccaggagatg 466
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 122 GTAACGATGTTGTCAAATTTAGCTGATTTTGGTTTTCAGAGAGTTGTCAAATACAAATGAT 181
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 467 cctacaccgattatgtagctacgagatggtaccgagctcctgaacttcttgaggagata 526
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 182 TATATACAGATTATGTTGCAACTCGTTGGTATAGATGCCCTGAATTATAGTTGGTGATA 241
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 527 ctccagtatggttcttcagtcgatatatatgggctattgggttgctttttgcagagcctcctga 586
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 242 TTGAATATGGGGTTAGTGTGTGATATNNGGCAATTTGGATGTGTATTAGCTGAAATGGTTA 301
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 587 caggccagccactgtgacctggaaatcagatgtggaccacactttatctgataaatcagaa 646
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 302 GTGGTGATGCAATATGGCCTGGTAGAGTGATGTTGATNAACTTTTTTTTGATATTTAAA 361
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 647 cactaggaataataatcccaaga 669
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 362 CATGGGTTCACTTCTGCCAAGA 384
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Search completed: May 2, 2002, 17:06:45
Job time: 11057 sec
```

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 17:09:45 ; Search time 116.84 Seconds
(without alignments)
1831.750 Million cell updates/sec

Title: US-09-671-050-11
Perfect score: 945
Sequence: 1 atggaagtgatgaaaaatt.....aggctactccgctcaaaagt 945

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*

5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*

6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	171.8	18.2	903	2	US-08-874-347-9 Sequence 9, Appli
2	171.8	18.2	903	3	US-09-093-522-9 Sequence 9, Appli
3	163.4	17.3	1476	2	US-08-969-106-1 Sequence 1, Appli
4	134.6	14.2	1089	1	US-08-154-915-1 Sequence 1, Appli
5	134.6	14.2	1089	2	US-08-464-517-37 Sequence 37, Appl
6	134.6	14.2	1089	2	US-08-246-361A-37 Sequence 37, Appl
7	134.6	14.2	1089	3	US-08-463-772-37 Sequence 37, Appl
8	134.6	14.2	1089	5	PCT-US93-09945-1 Sequence 1, Appli
9	131	13.9	1002	1	US-08-463-090B-3 Sequence 3, Appli
10	117.2	12.4	1070	1	US-08-463-090B-5 Sequence 5, Appli
11	103	10.9	2747	2	US-08-874-347-1 Sequence 1, Appli
12	103	10.9	2747	3	US-09-093-522-1 Sequence 1, Appli
13	98	10.4	1673	4	US-09-347-801-15 Sequence 15, Appl
14	97.8	10.3	1100	2	US-08-950-449A-18 Sequence 18, Appl
15	97.8	10.3	3774	2	US-08-950-449A-11 Sequence 11, Appl
16	97.8	10.3	3813	1	US-08-469-421-11 Sequence 11, Appl
17	97.8	10.3	3813	1	US-08-250-975-11 Sequence 11, Appl
18	97.8	10.3	3813	2	US-08-605-002A-11 Sequence 11, Appl
19	97.8	10.3	3813	5	PCT-US94-10529-11 Sequence 11, Appl
20	88.2	9.3	1381	2	US-08-950-449A-13 Sequence 13, Appl
21	88.2	9.3	1423	1	US-08-469-421-13 Sequence 13, Appl
22	88.2	9.3	1423	1	US-08-250-975-13 Sequence 13, Appl
23	88.2	9.3	1423	2	US-08-605-002A-13 Sequence 13, Appl
24	88.2	9.3	1423	5	PCT-US94-10529-13 Sequence 13, Appl
25	88.2	9.3	1539	3	US-09-286-904-1 Sequence 1, Appli
26	86.6	9.2	2624	1	US-08-032-382B-1 Sequence 1, Appli
27	85.4	9.0	1200	1	US-08-011-398B-3 Sequence 3, Appli

28	85.4	9.0	1200	1	US-08-464-051-3 Sequence 3, Appli
29	85.4	9.0	1200	2	US-08-462-498-3 Sequence 3, Appli
30	85	9.0	1429	4	US-09-347-801-13 Sequence 13, Appl
31	84.2	8.9	4453	1	US-08-770-761A-4 Sequence 4, Appli
32	84.2	8.9	4540	1	US-08-770-761A-6 Sequence 6, Appli
33	84.2	8.9	4621	1	US-08-770-761A-1 Sequence 1, Appli
34	83.8	8.9	1023	3	US-08-554-385-4 Sequence 4, Appli
35	83.2	8.8	912	2	US-08-627-610-13 Sequence 13, Appl
36	83.2	8.8	1424	1	US-08-403-634-3 Sequence 3, Appli
37	83.2	8.8	1424	4	US-08-913-441B-3 Sequence 3, Appli
38	82.2	8.7	1814	4	US-09-347-801-9 Sequence 9, Appli
39	81.6	8.6	2248	1	US-08-369-780-1 Sequence 1, Appli
40	81.6	8.6	2248	1	US-08-475-682-1 Sequence 1, Appli
41	81.6	8.6	2248	1	US-08-780-833-1 Sequence 1, Appli
42	81.6	8.6	2248	1	US-08-636-036-1 Sequence 1, Appli
43	81.6	8.6	2248	3	US-08-918-509-1 Sequence 1, Appli
44	81.6	8.6	2248	3	US-09-108-262-1 Sequence 1, Appli
45	75.4	8.0	1544	2	US-08-837-593-1 Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-08-874-347-9

; Sequence 9, Application US/08874347

; Patent No. 5863741

; GENERAL INFORMATION:

; APPLICANT: Limper, Andrew H.

; APPLICANT: Leof, Edward B.

; APPLICANT: Thomas, Charles F.

; APPLICANT: Gustafson, Michael P.

; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS

; TITLE OF INVENTION: CARINII

; NUMBER OF SEQUENCES: 26

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C., P.A.

; STREET: 60 South Sixth Street, Suite 3300

; CITY: Minneapolis

; STATE: MN

; COUNTRY: USA

; ZIP: 55402

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/874,347

; FILING DATE: 13-JUN-1997

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Ellinger, Mark S.

; REGISTRATION NUMBER: 34,812

; REFERENCE/DOCKET NUMBER: 07039/055001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 612-335-5070

; TELEFAX: 612-288-9696

; TELEX:

; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 903 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: Coding Sequence

; LOCATION: 1...900

; OTHER INFORMATION:

US-08-874-347-9


```
Query Match      18.2%; Score 171.8; DB 2; Length 903;
Best Local Similarity 52.3%; Pred. No. 2.4e-39;
Matches 458; Conservative 0; Mismatches 402; Indels 15; Gaps 3;

QY 1 atggaaaagtatgaaaaattagctaagactggagaagggctttatggggttgattcaaa 60
Db 1 ATGGAGCAATATCAGAGGTTAGAGAAGATTGGAGAAGGAACCTTATGGAGTTGTTATAAA 60
QY 61 tgcagaaaaacaaacctctggaacagtagtagctgtttaaaaaatttgggaatctgaagat 120
Db 61 GCAAAGGATCTTGAAAGTGGTACAATTGTAGCTCTTAAGAAAATCCGATTAGAAGCAGAA 120
QY 121 gatcctgtgtttaagaaaaatagcactaagagaaatcgtatgttgaagcaattaaaaacat 180
Db 121 GATGAGGGAGTTCCTAGTACAGCAATTCGTGAGATATCATTTCGAAAAGAGATGCACAAT 180
QY 181 ccaaatctgtgaacctcatcgagggtgttcaggagagaaaaaggaaatgcattagtttt 240
Db 181 GATAATGTTGTAAGACTTTTGAATATATAATTCATCAAGAGTCACGTTTATATCTTGTTTT 240
QY 241 gaatactgtgatcatacacacttttaaatgagctggaagaaacccaaatg-----gagtt 294
Db 241 GAATTTCTTGATCTGATTTAAAAAAGTATATGAATAGTATTCGAAAAGGACATGATGCTT 300
QY 295 gctgatggagtgcatacaaaagcgtattatggcgaacacttccaaactcttaattctgtcat 354
Db 301 GGTGCAGAAATGATTAAAAAGTTTATGTCACAACCTGTATCAGGTGTTAAATATTGTCAAT 360
QY 355 atacataactgtattcacagagatataaaaacctgaaaaatattctaataactaagcaagga 414
Db 361 TCTCATCGTATTCTTCATCGTACTTGAAACCCACAAAATCTTTATAGATCGAGAAGGA 420
QY 415 ataatacaagatttgtgacttcgggtttgca---caaatctgattccaggagatgcctac 471
Db 421 AATCTTAAATTAGCAGATTTTGGGCTTGCAAGGCGTTTGGTGTTCATTCGCTGCTTAT 480
QY 472 accgattatgtactacgagatggtaccgagctcctgaactcttctgtgggagatactcag 531
Db 481 ACTCATGAAGTTGTTACACTTTGGTATCGTCTCCAGAAGTCTTTTAGGTGGTCGACAA 540
QY 532 tatggttcttcagtcgatatatatgggctattggttggtttttgcagagctcctgacaggc 591
Db 541 TATGCAACACGCGCTTGATATATGAGACATTGGATGTATTTTGCAGAAATGGCTACAAAA 600
QY 592 cagccactgtggcctggaaaaatcagatgtggaccaactttatctgataaatcagaacacta 651
Db 601 AAGCCATTATTTCAGGTGATTCTGAAATTGATGAATAATTTAGAATATTTAGAATATTA 660
QY 652 ggaataattaatcccaagacatacaatcaatctttaaagagtaacgggttttccatggcatc 711
Db 661 GGGACTCCAGATGAAAAATTCCTGGCCCTGGTATTACATCTTATCCGGATTTTAAGGCAACT 720
QY 712 agtatacttgagccagaagacatggaactcttgagaaaaagttctcagatgttcatcct 771
Db 721 TTTCCAAAATGGTCACCA-----AAAAATCTTGGAGAATTAAATTACAGAACTTGATAGT 774
QY 772 gtggctctgaacttcatgaagggggtgtctgaaagatgaatccagatgacagattaaacctgt 831
Db 775 GATGGAATAGATTATTACAGAAAATGCTTAGATATATTATCCTGCTGAACGTATTAGCGCT 834
QY 832 tcccaactcctggagagctcctactttgattcttt 866
Db 835 AAAAAAGCTCTCGATCATCTTATTATTGATGATT 869
```

RESULT 2
US-09-093-522-9
; Sequence 9, Application US/09093522
; Patent No. 6015700
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.

```
APPLICANT: Gustafson, Michael P.  
TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C., P.A.  
STREET: 60 South Sixth Street, Suite 3300  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
APPLICATION DATA:  
APPLICATION NUMBER: US/09/093.522  
FILING DATE: 08-JUN-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/874,347  
FILING DATE: 13-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ellinger, Mark S.  
REGISTRATION NUMBER: 34,812  
REFERENCE/DOCKET NUMBER: 07039/055002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-335-5070  
TELEFAX: 612-288-9696  
TELEX:  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 903 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 1...900  
OTHER INFORMATION:  
US-09-093-522-9
```

```
Query Match      18.2%; Score 171.8; DB 3; Length 903;
Best Local Similarity 52.3%; Pred. No. 2.4e-39;
Matches 458; Conservative 0; Mismatches 402; Indels 15; Gaps 3;

QY 1 atggaaaagtatgaaaaattagctaagactggagaagggcttctatggggttgattcaaa 60
Db 1 ATGGAGCAATATCAGAGGTTAGAGAAGATTGGAGAAGGAACCTTATGGAGTTGTTATAAA 60
QY 61 tgcagaaaaacaaacctctggaacagtagtagctgtttaaaaaatttgggaatctgaagat 120
Db 61 GCAAAGGATCTTGAAAGTGGTACAATTGTAGCTCTTAAGAAAATCCGATTAGAAGCAGAA 120
QY 121 gatcctgtgtttaagaaaaatagcactaagagaaaaatacgtatgttgaagcaattaaaaacat 180
Db 121 GATGAGGGAGTTCCTAGTACAGCAATTCGTGAGATATCATTTCGAAAAGAGATGCACAAT 180
QY 181 ccaaatcttgtgaacctcatcgagggtgttcaggagagaaaaaggaaatgcattagtttt 240
Db 181 GATAATGTTGTAAGACTTTTGAATATATAATTCATCAAGAGTCACGTTTATATCTTGTTTT 240
QY 241 gaatactgtgatcatacacacttttaaatgagctggaagaaaaacccaaatg-----gagtt 294
Db 241 GAATTTCTTGATCTTGATTTAAAAAAGTATATGAATAGTATTCGAAAAGGACATGATGCTT 300
QY 295 gctgatggagtgcatacaaaagcgtattatggcaaacacttcaagctcttaattctgtcat 354
Db 301 GGTGCAGAAATGATTAAAAAGTTTATGTCACAACCTGTATCAGGTGTTAAATATTGTCAAT 360
```



```

;
; TITLE OF INVENTION: Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/154,915
; FILING DATE: 19-NOV-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/991,997
; FILING DATE: 17-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-026
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1089 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 13..888
; US-08-154-915-1

Query Match 14.2%; Score 134.6; DB 1; Length 1089;
Best Local Similarity 52.3%; Pred No. 8.5e-29;
Matches 346; Conservative 0; Mismatches 309; Indels 6; Gaps 2;

QY 1 atggaaaagtatgaaaaattagctaagactggagaagggtcttatggggtgtattcaaa 60
Db ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
13 ATGCAGAAATACGAGAACTGGAAAAGATTGGGGAAGGCACCTACGGAAGTGTGTTCAAG 72
QY 61 tgcagaaacaaacctctggaacaagtagtagctgttaaaaaatttgtggaatctgaagat 120
Db || | || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
73 GCCAAAAACCGGAGACTCATGAGATCGTGGCTCTAAACCGGTGAGGCTGCATGACGAT 132
QY 121 gatcctgtgttaagaaaaatagcactaagagaaatacgtatgttgaaagcaattaaacat 180
Db || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
133 GATGAGGGTGTGCCGAGTTCGCCCTCCGGGAGATCTGCCTACTCAAGGAGCTGAAGCAC 192
QY 181 ccaaatcttgtgaacctcatcgagggtgttcaggagaaaaaggaaatgcatttattttt 240
Db || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
193 AAGAACATCGTCAGGCTTCATGACGTCCTGCACAGCGACAAGAGCTGACTTTGGTTTTT 252
QY 241 gaatactgtgatcatacaccttttaaatgagctggaaagaaaccccaatggagttgctgat 300
Db || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
253 GAATTCTGTGACCAGGACCTGAAGAAGTATTTTGACAGTTGCAATGGTGACCTCGATCCT 312
```

```

QY 301 ggagtgatcaaaagcgtattatgggcaaacacttcaagctcttaattctgtcatatacat 360
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
313 GAGATTGTAAGTCATTCCTCTTCCAGCTACTAAAAGGGCTGGGATTCTGTATAGCCGC 372
QY 361 aactgtattcacagagagataaaaaacctgaaatatcttaataaagcaaggaataatc 420
Db || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
373 AATGTGCTACACAGGGACCTGAAGCCCCCAGAACCTGTCTAATAACAGGAATGGGGAGCTG 432
QY 421 aagatttgtacttcgggtttgcacaaa---ttctgattccaggagatgcctacacccgat 477
Db || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
433 AAATGGCTGATTTTGGCCTGGCTCGAGCCTTTGGGATTCCCGTCCGCTGTACTCAGCT 492
QY 478 tatgtactacgagatggtaccgagctcctgaaacttcttgtggagatactcagtatggt 537
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
493 GAGGTGGTCACACTGTGTTACCGCCCGGATGTCCTCTTTGGGGCCAAGCTGTACTCC 552
QY 538 tcttcagtcgatatatgggtattgtgtgttttgcagagct---cctgacagggccag 594
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
553 ACGTCCATCGACATGTGCTCAGCCGGCTGCATCTTTGCAGAGCTGGCCAATGCTGGGCGG 612
QY 595 ccactgtgcctggaaaatcagatgtggaccacatttatctgataatacagaacactagga 654
Db || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
613 CCTCTTTTCCCGGCAATGATGTCATGACCAGTTGAAGAGGATCTTCCGACTGCTGGGG 672
QY 655 a 655
Db |
673 A 673
```

```

RESULT 5
US-08-464-517-37
; Sequence 37, Application US/08464517
; Patent No. 5869640
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,517
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-004C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1089 base pairs
; TYPE: nucleic acid
```

```

; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 13..888
US-08-464-517-37

Query Match      14.2%; Score 134.6; DB 2; Length 1089;
Best Local Similarity 52.3%; Pred. No. 8.5e-29;
Matches 346; Conservative 0; Mismatches 309; Indels 6; Gaps 2;

QY 1 atggaaaagtatgaaaaattagctaaagtagtagctggagaaggtcttattggggtgtattcaaa 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13 ATGCAGAAATACGAGAAACTGGAAAAGATTGGGAAGGCACCTACGGAACCTGTGTTCAAG 72

QY 61 tgcagaaacaaacctctggacaagttagtagctgtttaaaaaatttgtggaatctgaagat 120
   || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73 GCCAAAAACCGGAGACTCATGAGATCGTGGCTCTAAACGGGTGAGGCTGGATGACGAT 132

QY 121 gatcctgtgttaagaaaaatagcactaagagaaaaatcgtatgttgaaagcaattaaaaacat 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 133 GATGAGGSGTGCAGGCTTCATGACGTCCTCCGCCCTCCGAGAGATCTGCCTACTCAAGGAGCTGAAGCAC 192

QY 181 ccaaatcttgtgaacctcatcgagggtgttcaggagaaaaaggaaaaatgcattagtttt 240
   || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 193 AAGAATCGTCAGGCTTCATGACGTCCTGCACAGCGACAAGAGCTGACCTTGGTGT 252

QY 241 gaactgtgtatcaaaaaagcgtattatggcaaacacttcaagctcttaattctgtcatatacat 360
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 313 GAGATTGTAAGTCATCTCTCCAGCTACTAAAAGSGCTGGGATTCTGTCTATAGCCGC 372

QY 361 aactgtatlcacagagatataaaaacctgaaaaatttcttaataactaagcaaggaataatc 420
   || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 373 AATGTGCTACACAGGGACCTGAAGCCCCAGAACCTGCTAATAAACAGGAATGGGGAGCTG 432

QY 421 aagatttgtgacttcgggtttgcacaaa---ttctgattccaggagatgcctacacogat 477
   || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 433 AAATTGGCTGATTTTGGCCTGGCTCGAGCCTTTGGGATTCCCGTCCGCTGTACTCAGCT 492

QY 478 tatgtagctacgagatggtaccgagctcctgaaccttcttggggagatactcagtatggt 537
   || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 493 GAGGTGTCACACTGTGTTACCGCCACCGGATGTCCTCTTTGGGGCCAAAGCTGTACTCC 552

QY 538 tcttcagtcgatatatgggctattggttgtgttttgcagagct---cctgacagggccag 594
   || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 553 ACGTCCATCGACATGTGGTCAGCGGCTGCATCTTTGCAGAGCTGGCCAAATGCTGGCGG 612

QY 595 ccaactgtggccctggaaaaatcagatgtggaccaactttatctgtataaatcagaacactagga 654
   || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 613 CCTCTTTTCCCGGCAATGATGTCGATGACCAAGTTGAAGAGGATCTTCCGACTGCTGGGG 672

QY 655 a 655
Db 673 A 673

RESULT 6
US-08-246-361A-37
; Sequence 37, Application US/08246361A
; Patent No. 598582
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
```

```

; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/246,361A
; FILING DATE: 19-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-004C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1089 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 13..888
US-08-246-361A-37
```

```

Query Match      14.2%; Score 134.6; DB 2; Length 1089;
Best Local Similarity 52.3%; Pred. No. 8.5e-29;
Matches 346; Conservative 0; Mismatches 309; Indels 6; Gaps 2;

QY 1 atggaaaagtatgaaaaattagctaaagtagtagctggagaaggtcttattggggtgtattcaaa 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13 ATGCAGAAATACGAGAAACTGGAAAAGATTGGGAAGGCACCTACGGAACCTGTGTTCAAG 72

QY 61 tgcagaaacaaacctctggacaagttagtagctgtttaaaaaatttgtggaatctgaagat 120
   || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73 GCCAAAAACCGGAGACTCATGAGATCGTGGCTCTAAACGGGTGAGGCTGGATGACGAT 132

QY 121 gatcctgtgttaagaaaaatagcactaagagaaaaatcgtatgttgaaagcaattaaaaacat 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 133 GATGAGGSGTGCAGGAGTTCCGCCCTCCGCGAGATCTGCCTACTCAAGGAGCTGAAGCAC 192

QY 181 ccaaatcttgtgaacctcatcgagggtgttcaggagaaaaaggaaaaatgcattagtttt 240
   || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 193 AAGAATCGTCAGGCTTCATGACGTCCTGCACAGCGACAAGAGCTGACCTTGGTGT 252

QY 241 gaatactgtgatcacacttttataatgagctggaaaaaggaaaaatgagttgctgtgat 300
   |||| |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 253 GAATTCTGTGACCGGACCTGAAGAAGTATTTGACAGTTGCAATGGTGACCTGCATCCT 312

QY 301 ggagtgtatcaaaaaagcgtattatggcaaacacttcaagctcttaattctgtcatatacat 360
   || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 313 GAGATTGTAAGTCATCTCTCCAGCTACTAAAAGSGCTGGGATTCTGTCTATAGCCGC 372

QY 361 aactgtatlcacagagatataaaaacctgaaaaatttcttaataactaagcaaggaataatc 420
   || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 373 AATGTGCTACACAGGGACCTGAAGCCCCAGAACCTGCTAATAAACAGGAATGGGGAGCTG 432
```

QY	421	aagatttgacttcgggtttgcacaaa---ttctgattccaggagatgccatacccgat	477
Dd	433	AAATTGGCTGATTTTGGCCTGGCTCGAGCCCTTTGGGATTCCCGTCCCGTGTTACTCAGCT	492
QY	478	tatgtagctacgagatggttaaccgagctcctgaaccttcttgtgggagatactcagtatggt	537
Dd	493	GAGGTGGTCACACTGTGGTACC GCCCACCGGATGTCTCTTTGGGGCAAGCTGTACTCC	552
QY	538	tcttcagtcgatatatgggctattggttgtgtttttgcagagct---octgacaggcccag	594
Dd	553	ACGTCCATCGACATGTGGTACGCCGCTGCATCTTTGCAGAGCTGGCCAATGCTGGGCGG	612
QY	595	ccactigtggccctgggaaaaatcagatgtggaccaactttatctgataatcagaacactagga	654
Dd	613	CCTCTTTTTCCCGGCAATCATGTCTCATGACCAGTGAAGAGGATCTTCCGACTGCTGGGG	672
QY	655	a 655	
Dd	673	A 673	

RESULT 7
US-08-463-772-37
; Sequence 37, Application US/08463772
; Patent No. 6066501
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50

```

Query Match          14.2%;   Score 134.6;   DB 3;   Length 1089;
Best Local Similarity 52.3%;   Pred. No. 8.5e-29;
Matches 346;   Conservative 0;   Mismatches 309;   Indels 6;   Gaps 2;

QY 1 atggaagaagtatgaaaaattagctaaagactggagaaggtctttatgggggttgattcctc 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13 ATGCAGAAATACGAGAAACTGGAAGAGATTGGGAAGGCACCTACGGAACCTGTTGTTCAAG 72

QY 61 tgcagaacaaaaacctctggcaagtagtagctgtttaaaaaatttgtggaatctgaaagat 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73 GCCAAAACCGGAGACTCATGAGATCGTGGCTCTAAAACGGGTGAGGCTGGATGACCAT 132

QY 121 gatcctgttggtaagaaaaatagcactaagagaaatacgtatgttgaaagcaattaaaaacat 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 133 GATGAGGGTGTGCCGAGTTCCGCCCTCCGGGAGATGTCCTACTCAAGGAGCTGAAGCAC 192

QY 181 ccaaatcttgtgaacctcatcgaggtgttcaggagaaaaaggaaaaatgcatttagttttt 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 193 AAGAACATCGTCAGGCTTCATGACGTCTCCAGCTACTAAAAGGGCTGGGATTTGGT 252

QY 241 gaatactgtgatcatcacacttttaaatgagctggaagaaaccccaaatgagttgctgat 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 253 GAATTCTGTGACCAGGACCTGAAGAAGTATTTGACAGTTGCAATGGTGACCTCGATCCT 312

QY 301 ggaagtatcaaaagcgtattatggcaaacacattcaagctcttaatttctgtcatatacat 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 313 GAGATTGTAAAGTCATTCTCTTCCAGCTACTAAAAGGGCTGGGATTTCTGTATAGCCGC 372

QY 361 aactgtattcacagagatataaaacctgaaaattattcttaataactaaaggaataaatc 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 373 AATGTGCTACACAGGACCTGAAGCCCCAGAACCTGCTAATAACAGGAATGGGAGCTG 432

QY 421 aagatttgtgacttcgggtttgcaaaaa---ttctgattccaggagatgcctacacccgat 477
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 433 AAATTTGGCTGATTTTGGCTGGCTCGAGCCTTTGGGATTTCCCGTCCGCTGTTACTCAGCT 492

QY 478 tatgtagctacagataggtaccagagctcctgaactctctgtgtggagatactcagtatggt 537
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 493 GAGGTGGTCACACTGTGGTACCGCGGCTGCATCTTTGGGGCCAAAGCTGACTCC 552

QY 538 tcttcagtcgatataatgggtctatbggtgtgtgtttttgtcagagct---cctgacagggccag 594
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 553 ACGTCCATCGACATGTGGTCAGCGGGCTGCATCTTTGCAGAGCTGGCCCAATGCTGGGGCG 612

QY 595 ccactgtgacctggaataatcagatgtggaccaactttatctgtataatcagaaactagga 654
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 613 CCTCTTTTCCCGGCAATGATGCGATGACCAAGTTGAAGAGGATCTTCCGACTGCTGGGG 672

QY 655 a 655
    |
Db 673 A 673

RESULT 8
PCT-US93-09945-1
; Sequence 1, Application PC/TUS9309945
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Cyclin Complex Rearrangement and Uses Related
; TITLE OF INVENTION: Thereto
; NUMBER OF SEQUENCES: 4
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/09945
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992

```


Db 419 CTCATCGAGTTTACATCGTGATTAAACCACAAAAATTATTGATTGATAAAGRAGGGA 478

QY 416 taatcaagatttgtgaotttcgggtttgcacaaa---ttctgattccaggagatgcctaca 472

Db 479 ATTTAAAAATAGCAGATTTTGGATTAGCTCGAGCATTTGGAGTTCCATTAAAGAGCATATA 538

QY 473 ccgattatgtactacgagatggtaccgagctcctgaacttcttggggagatactcagt 532

Db 539 CTCATGAAGTTGTCACTTTATGGTATCGAGCTCCGAAATCTTGTAGGAGGGAACAAT 598

QY 533 atggttcttcagtcgatatatagggtctattggtgtgttttgcagagctcctgacaggcc 592

Db 599 ATTCCACTGGGGTAGATATGTGGTCTGTGGATGTATATTTGCTGAAATGCTGTAATAGGA 658

QY 593 agccactgtggcctggaaaaatcagatgtggaccacatttatctgtataatcagaacactag 652

Db 659 AACCAATTATTTCCCTGGTGATTTCAGAAATTGATGAAATTTTCCGAATTTTCCGAAATTTTAG 718

QY 653 gaaaa 657

Db 719 GAACA 723

RESULT 10

US-08-463-090B-5

; Sequence 5, Application US/08463090B

; Patent No. 5801015

; GENERAL INFORMATION:

; APPLICANT: Cottarel, Guillaume

; APPLICANT: Damagnez, Veronique

; APPLICANT: Draetta, Guilio

; TITLE OF INVENTION: Cell-Cycle Regulatory Proteins from

; TITLE OF INVENTION: Human Pathogens, and Uses Related Thereto

; NUMBER OF SEQUENCES: 25

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley, Hoag & Eliot, LLP

; STREET: One Post Office Square

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII (text)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/463,090B

; FILING DATE: 05-JUN-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Vincent, Matthew P.

; REGISTRATION NUMBER: 36,709

; REFERENCE/DOCKET NUMBER: MIV032.01

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 832-1299

; TELEFAX: (617) 832-7000

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1070 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: both

; TOPOLOGY: both

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 30..1058

US-08-463-090B-5

Query Match 12.4%; Score 117.2; DB 1; Length 1070;

Best Local Similarity 49.7%; Pred. No. 6.9e-24;

Matches 327; Conservative 0; Mismatches 328; Indels 3; Gaps 1;

QY 7 aagtatgaaaaattagctaagactgggagaaaggtccttatgggggttgatttcaaatgcaga 66

Db 96 AATTACACCAAGAAAAAGTAGGGGAGGTACATATGCTGTTGTGTACTTTGGGTAAA 155

QY 67 aacaaaaacctctggacaagttagtagctgttlaaaaaatttggtaactctgaagatgacct 126

Db 156 CAAATCTCCACCACCAACGTCAAATTGCCATCAAAAGAAATCAAAACAGGATTATTCAAAGAT 215

QY 127 gttgttaagaaaaatagcactaagagaaaatcacgtatgttgaagcaattlaaaacatccaaat 186

Db 216 GGGTTGGATATGTCAGCATTTGAGAGAAGTCAAAATATTTGCAAGAATTGAAACATCCCAAT 275

QY 187 cttgtgaacctcatcgagggtgttcaggagagaaaaagcaaatgcatttagttttgaatac 246

Db 276 GTTATTGAACCTAGTAGATGTATTTTCAGCAACAAATAAATTAAATTTGGTATTAGAATTT 335

QY 247 tgtgatcatcacacttttaaatgagctggaaagaaaccccaatggagtgtgctgatggagtg 306

Db 336 CTACCTTGGGATTTGGAAGTGTGTGATCAAGATAAATCGATTGTTTCAATCAGCAGAT 395

QY 307 atcaaaaagcgtattatggcaaacacacttcaagctcttaatttctcatatatacataaactgt 366

Db 396 ATCAAAATCATGGCTTTTAAATGACATTACGTGGGATACATCATTTGTATCGGAATTTTATT 455

QY 367 attcacagagatatataaaacctgaaaaatattctataataagcaaggaataaatcaagatt 426

Db 456 TTACATCGTGATTTGAAACCAAAATAATTTATTATTGGCACCCGGATGGACAATTGAAATA 515

QY 427 tgtgacttcgggtttgcacaaattctg---attccaggagatgcctacaccgattatgta 483

Db 516 GCGGATTTTGGTCTTGCACGAGCTTTGGTAAATCCTAATCAAGATTTTATCATCTAATGTT 575

QY 484 gctacgagatggtaccgagctcctgaacttcttctgtgggagatactcagtatggttttca 543

Db 576 GTCACTAGATGGTATAGAGCCCTGAATTATTATTGGTCTCGACATTACACTGGAGCA 635

QY 544 gtcgatataatggctattgtgtgttttgcagagctcctgacagggccagccactgtgg 603

Db 636 GTTGATATCTGGTCAATAGGTATAATATTGCTGAATTAATGCTTCGAATACCTTATTG 695

QY 604 cctggaaaaatcagatgtggaccacatttatctgtataatcagaacactaggaaaaataa 661

Db 696 CCAGGTAAGATGACGTTGATCAATTAGATTAGATTATGAGCTTATGGGACACCAA 753

RESULT 11

US-08-874-347-1

; Sequence 1, Application US/08874347

; Patent No. 5863741

; GENERAL INFORMATION:

; APPLICANT: Limper, Andrew H.

; APPLICANT: Leof, Edward B.

; APPLICANT: Thomas, Charles F.

; APPLICANT: Gustafson, Michael P.

; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS

; TITLE OF INVENTION: CARINII

; NUMBER OF SEQUENCES: 26

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C., P.A.

; STREET: 60 South Sixth Street, Suite 3300

; CITY: Minneapolis

; STATE: MN

; COUNTRY: USA

; ZIP: 55402

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/874,347

; FILING DATE: 13-JUN-1997

Db 1351 TTGGAGCTGAAATGATCAAAAAATTTATGTCTCAACTTGTTATCAGGTGTTAAATATTGTC 1410
QY 353 atatacataactgtattcacagagatatataaacctgaaatatattctaataactaagcaag 412
Db 1411 ATTCTCATCGTATTCTTCATCGTGACTTGAGGCCCCAAAATCTTCTATTATGATCGAGAAG 1470
QY 413 gaataatcaagatttgtgacttcgggtttgac---aaattctgattccaggagatgcct 469
Db 1471 GAAATCTTAAACTTGCTGATTTCGGGCTTGCTCGGCATTTGGTGTTCCTTTACGTGGTT 1530
QY 470 acaccgattatgtagctacgagatggtaccagagctcctgaactcttctgtgggagatactc 529
Db 1531 ATACTCATGAGGTGTTACACTTTGGTATCGTGCTCCAGAAGTCTTTTAGGTGGTCGAC 1590
QY 530 agtatggttcttcagtcgatatatatgggctattgtgtgtgtttttgcagagotcctgcagag 589
Db 1591 AATATGCAACAGCACTTGATATTGGAGCATTTGGATGTATTTTGTCTGAAATGGCTACAA 1650
QY 590 gccagccactgtggcctggaataatcacgatgtggaccaactttatctgtataatcagaaca 648
Db 1651 AAAAACCGTTATTTCAGGCGGATTCTGAAAATTGATGAAATATTCAGAAATATTAGGTCA 1709
RESULT 13
US-09-347-801-15
; Sequence 15, Application US/09347801
; Patent No. 6262345
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Lee, Jian Ming
; TITLE OF INVENTION: Plant Protein Kinases
; FILE REFERENCE: BB-1171
; CURRENT APPLICATION NUMBER: US/09/347,801
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: 60/092,438
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 1673
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1349)
US-09-347-801-15

Query Match 10.4%; Score 98; DB 4; Length 1673;
Best Local Similarity 60.4%; Pred. No. 2.3e-18;
Matches 180; Conservative 0; Mismatches 115; Indels 3; Gaps 1;
QY 370 cacagagatataaaacctgaaaaatttctaataactaagcaa---ggaataatcaagatt 426
Db 599 cacagggatgtgaaccacaaaaatgtttgttgatcctctaaccatcaagtcgaagtc 658
QY 427 tgtgacttcgggtttgcacaaattctgattccaggagatgcctacacgcgattatgtagct 486
Db 659 tgtgactttggaagtgcaaaagtctctgttacctggtgaccccaacatagcatacatatgc 718
QY 487 acgagatggtaccgagctcctgaactcttgtgtggagatactcagtatggtttctcagtc 546
Db 719 tctcgctactatcgtgtcctcctgagctcatatttgtgcaactgaatatacaactccaata 778
QY 547 gatatatgggctattggttgtttttgcagagctcctgcagcaggccagccactgtggcct 606
Db 779 gacatatggtcagctggatgtgttcttgcagagctacttcttggtcagcctctgtttcca 838
QY 607 ggaataatcagatgtggaccacactttatctgataatcagaacactaggaaaaattaatcc 664
Db 839 ggagagactcgggttgatcagctagtgaggattatcaaggttcttgggtactccaacc 896

RESULT 14
US-08-950-449A-18
; Sequence 18, Application US/08950449A
; Patent No. 5955366
; GENERAL INFORMATION:
; APPLICANT: Lee, John C.
; APPLICANT: Adams, Jerry L.
; APPLICANT: Gallagher, Timothy F.
; APPLICANT: Green, David W.
; APPLICANT: Heys, J. Richard
; APPLICANT: McDonnell, Peter
; APPLICANT: McNulty, Dean E.
; APPLICANT: Strickler, James E.
; APPLICANT: Young, Peter R.
; TITLE OF INVENTION: Drug Binding Protein
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: Corporate Intellectual Property/P.O. Box
; STREET: 1539
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/950,449A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/123,175
; FILING DATE: 17-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hecht, Elizabeth J.
; REGISTRATION NUMBER: P-41,824
; REFERENCE/DOCKET NUMBER: P50195-1D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 270-5009
; TELEFAX: (610) 270-5090
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1100 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-950-449A-18

Query Match 10.3%; Score 97.8; DB 2; Length 1100;
Best Local Similarity 46.5%; Pred. No. 2.1e-18;
Matches 404; Conservative 0; Mismatches 447; Indels 18; Gaps 2;
QY 37 gggctcttatggggttgattccaatgcagaaaaaacctctggacaagttagtctgtt 96
Db 108 GCGCGCTATGGCTCTGTGTGTGCTCTTTTGACACAAAACGGGTTACGTGTGGCAGTG 167
QY 97 aaaaaatttgtggaatctgaagatgatcctgttgttaagaaaaatagcactaagagaaata 156
Db 168 AAGAAAGCTCTCCAGACCATTTTCAGTCCATCATTCATGCGAAAAGAACCTACAGAGAACTG 227
QY 157 cgtatgttgaagcaattaaaaacatccaaatcttgtgaacctcatcgagggttcaggaga 216
Db 228 CGGTTACTTAAACATATGAACATGAAAATGTGATGTGCTCTGTTGGACGTTTTCACACCT 287
QY 217 aaaaggaataatgcatttagtttttgaatactgtgatcatacacttttaaatgagctgga 276

Db 288 GCAAGGTCCTGGAGGAATTCAATGATGTGTATCTGGTGACCCATCTCATGGGGCAGAT 347
QY 277 agaaaccaaagt-----gagttgctgatggagtgatcaaaaagcgtattatgg 324
Db 348 CTGAACAACATTGTGNAATGTCAGAAGCTTACAGATGACCATGTTTCAGTTCCTTATCTAC 407
QY 325 caaacacttcaagctcttaatttctgtcatatacataaactgtattcacagagataaaaa 384
Db 408 CAAATTCTCCGAGGTCTAAAGTATATACATTACAGCTGACATAATTACAGGGACCTAAAA 467
QY 385 cctgaaaaatattctaactaagcaaggaataatcaagatttgtgacttcgggtttgca 444
Db 468 CCTAGTAATCTAGCTGTGAATGAAGACTGTGAGCTGAAGATTCTGGATTTTGGACTGGC- 526
QY 445 caaattctgattccaggagatgcctacacccgattatgtagtacgagatggtaaccgagct 504
Db 527 -----TCGGCACACAGATGATGAATGACAGGCTACGTGGCCACTAGGTGGTACAGGGCT 581
QY 505 cctgaacttcttgtgggagatactcagtatggttcttcagtcgatataatgggctatttgtt 564
Db 582 CCTGAGATCATGCTGAACCTGGATGCATTACAACACAGACAGTTGATATTTGGTCAGTGGGA 641
QY 565 tgtgtttttgcagagctcctgacagccagccactgtggtcctggaaaaatcagatgtggac 624
Db 642 TGCATAATGGCCGAGCTGTGACTGGAAGAACAATTGTTTCCCTGGGTACAGACCATATTAAAC 701
QY 625 caactttatctgataatcagaacacactagggaaaaataatcccaagacatcaatcaatttt 684
Db 702 CAGCTTCAGCAGATTATGGCTCTGACAGGAACACACCCCGCTTATCTCATTAACAGGATG 761
QY 685 aaaagtaacgggtttttccatggccatcagtatatacctgagccagaagacatggaaaactctt 744
Db 762 CCAAGCCATGAGGCAAGAAACTATATTTCAGTCTTTGACTCAGATGCCGAAGATGAACTTT 821
QY 745 gaggaaaagtctcagatgttcacctgtggtcctggaacttcacatgaagggtgtctgaag 804
Db 822 GCGAATGTATTATTATTTGGTGCCAAATCCCGTGGCTGTGCGACTTGCTGGAGAGATGCTTGTA 881
QY 805 atgaatccagatgacagattaacctgttcccactcctgtggagagctcctactttgattct 864
Db 882 TTGGACTCAGATAAGAGAATTACAGCGGCCCAAGCCCTTGCACATGCCTACTTTGCTCAG 941
QY 865 ttccaagaggcccaaatcaaaagaaaaagc 893
Db 942 TACCACGATCCTGATGATGAACCAAGTGGC 970

RESULT 15
US-08-950-449A-11
; Sequence 11, Application US/08950449A
; Patent No. 595366
; GENERAL INFORMATION:
; APPLICANT: Lee, John C.
; APPLICANT: Adams, Jerry L.
; APPLICANT: Gallagher, Timothy F.
; APPLICANT: Green, David W.
; APPLICANT: Heys, J. Richard
; APPLICANT: McDonnell, Peter
; APPLICANT: McNulty, Dean E.
; APPLICANT: Strickler, James E.
; APPLICANT: Young, Peter R.
; TITLE OF INVENTION: Drug Binding Protein
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: Corporate Intellectual Property/P.O. Box
; STREET: 1539
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/950,449A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/123,175
; FILING DATE: 17-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hecht, Elizabeth J.
; REGISTRATION NUMBER: P-41,824
; REFERENCE/DOCKET NUMBER: P50195-1D1
; TELEPHONE: (610) 270-5009
; TELEFAX: (610) 270-5090
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3774 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; CELL TYPE: Monocyte
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 379..1461
; US-08-950-449A-11

Query Match 10.3%; Score 97.8; DB 2; Length 3774;
Best Local Similarity 46.5%; Pred. No. 3.7e-18;
Matches 404; Conservative 0; Mismatches 447; Indels 18; Gaps 2;

QY 37 gggtcttatggggttgattcaaaatgcagaaaaaaacctctggacaagtagtagctgtt 96
Db 459 GCGCCTATGGCTCTGTGTGCTGCTTTTGACACAAAAACGGGGTTACGTGTGGCAGTG 518
QY 97 aaaaaatttgtggaatctgaagatgatccctgttgaagaaaaatagcactaagagaaaa 156
Db 519 AAGAAGCTCTCCAGACCATTTCAGTCCCATTCATTCGCAAAAAGAACCTACAGAGAACTG 578
QY 157 cgtatgttgaagcaattaaaaacatcccaaatcttgtgaacctcatcgagggtgttcaggaga 216
Db 579 CGGTACTTAAACATATGAACATGAAATATGATTTGGTCTGTTGGACGTTTTTACACCT 638
QY 217 aaaaggaagaatgcatttagtttttgaatacttgtatcatcacacttttaaatgagctggaa 276
Db 639 GCAAGGTCTCTGGAGGAATTCAAATGATGTGTATCTGGTGACCCCATCTCATGGGGCAGAT 698
QY 277 agaaacccaaaatg-----gagttgctgatggagtgatcaaaaagcgtattatgg 324
Db 699 CTGAACAACATTGTGAAATGTCAGAAGCTTACAGATGACCATGTTTCAGTTCCTTATCTAC 758
QY 325 caaacacttcaagctcttaatttctgtcatatacataaactgtattcacagagataaaaa 384
Db 759 CAAATTCTCCGAGGTCTAAAGTATATACATTACAGCTGACATAAATTACAGGGACCTAAAA 818
QY 385 cctgaaaaatattctaactaagcaaggaataatcaagatttgtgacttcgggtttgca 444
Db 819 CCTAGTAATCTAGCTGTGAATGAAGACTGTGAGCTGAAGATTCTGGATTTTGGACTGGC- 877
QY 445 caaattctgattccaggagatgcctacacccgattatgtagtacgagatggtaaccgagct 504
Db 878 -----TCGGCACACAGATGATGAATGACAGGCTACGTGGCCACTAGGTGGTACAGGGCT 932
QY 505 cctgaacttcttgtgggagatactcagtatggttcttcagtcgatataatgggctatttgtt 564

Db	933	CCTGAGATCATGCTGAACTGGATGCATTACAA	CCAGACAGATTGATATTGGTCACTGGGA	992		
Qy	565	tgtgttttttg	cagagctccctgacagggccagccactgtg	gacctgga	aaatcagatgtggac	624
Db	993	TGCATAATG	GCCGAGCTGTTGACTGGAGAAGACATTGTT	CTGGTACAGACCATATTAAC	1052	
Qy	625	caactttatctgataatcagaacacactagga	aaattaatcccaagacatcaatcaatcttt	684		
Db	1053	CAGCTTCAGCAGATTATGCGTCTGACAGGA	ACACCCCCCGCTTATCTCATTAACAGGATG	1112		
Qy	685	aaaagtaacgggtttttcccatggcatcagta	tatacctgagccagagaacatgga	aaactctt	744	
Db	1113	CCAAGCCATGAGGCAAGAAACTATATTCA	CTCTTTGACTCAGATGCCGAAGATGA	AACTTT	1172	
Qy	745	gaggaagaagtctcagatgtttcatcctgtg	gctctgaacttc	atgaaggggtg	ctgaag	804
Db	1173	GCGAATGTATTTATTTGGTGCCCAATCCCC	CTGGCTGTCGACTTGC	TGGAGAGATGCTTTGA	1232	
Qy	805	atgaatccagatgacagattaacctgttcc	caactcctggagagagctc	actttgattct	864	
Db	1233	TTGGACTCAGATAAGAGAAATTACAGCGG	CCCCAAGCCCTTGCACATGCCTACTTTG	CTCAG	1292	
Qy	865	tttcaagaggcccaaat	taaaagaaagc	893		
Db	1293	TACCACGATCCTGATGATGA	ACCAGTGGC	1321		

Search completed: May 2, 2002, 17:10:02
Job time: 9674 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 18:43:31 ; Search time 170.24 Seconds
(without alignments)
5242.455 Million cell updates/sec

Title: US-09-671-050-9
Perfect score: 1041
Sequence: 1 atgaaaagtatgaaaaatt.....agggtactccgctcaaaagt 1041

Scoring table: OLIGO_MUC
Gapop 60.0 , Gapext 60.0
Searched: 930621 seqs, 428662619 residues

Word size : 20
Total number of hits satisfying chosen parameters: 781

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : N_Geneseq_1101.*
1: /SIDS1/gcgdata/geneseq/geneseqn/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1041	100.0	1041	22	Human kinase cDNA
2	1025	98.5	1068	22	Human kinase cDNA
3	652	62.6	945	22	Human kinase cDNA
4	652	62.6	972	22	Human kinase cDNA
5	572	54.9	1819	22	Human kinase cDNA
6	454	43.6	561	22	Human kinase cDNA
7	454	43.6	594	22	Human kinase cDNA
8	454	43.6	1083	22	Polynucleotide seq
9	368	35.4	911	22	Novel protein kina
c 10	40	3.8	12729	22	Human neuroblastom
11	33	3.2	105	21	Human secreted pro

33	12	3.2	411	14	AAQ61041	Human brain Expres
33	c 13	3.2	2286	22	AAF26586	DNA encoding human
33	c 14	3.2	5600	22	AAI60415	Human polynucleoti
33	c 15	3.2	6027	22	AAI58629	Human polynucleoti
31	c 16	3.0	32191	22	AAI64117	Human bladder rela
30	c 17	2.9	300	20	AAx98426	Human cancer cell
30	c 18	2.9	752	20	AAx98945	Human validated ca
30	c 19	2.9	1345	21	AAc93349	Human secreted pro
30	c 20	2.9	1718	22	AAH72939	Human cervical can
30	c 21	2.9	1762	22	AAS08146	Mammalian toxicolo
30	c 22	2.9	1804	22	AAH17495	Human cDNA sequenc
30	c 23	2.9	1916	22	AAH18014	Human cDNA sequenc
29	c 24	2.8	9439	22	AAI63723	Human kidney relat
29	c 25	2.8	237326	19	AAV57903	Hereditary haemoch
28	c 26	2.7	261	21	AAc05941	Human secreted pro
28	c 27	2.7	323	21	AAc05857	Human secreted pro
28	c 28	2.7	531	22	AAH13269	Human cDNA clone (
28	c 29	2.7	631	21	AAAI6349	Human cDNA clone (
28	c 30	2.7	805	22	AAH04221	Human GSHS related
28	c 31	2.7	1001	21	AAH51279	Human GLCL related
28	c 32	2.7	1001	21	AAH51559	Nucleotide sequenc
33	c 33	2.7	1469	22	AAf55366	Human transmembran
34	c 34	2.7	1562	21	AAZ56728	Human EST-derived
35	c 35	2.7	1624	22	AAH98793	Nucleotide sequenc
36	c 36	2.7	1763	22	AAf53365	Human cDNA sequenc
37	c 37	2.7	1763	22	AAf90192	Human secreted pro
38	c 38	2.7	1791	22	AAH18124	Human cDNA sequenc
39	c 39	2.7	1792	21	AAc59834	Human ORFX ORF2491
40	c 40	2.7	1850	22	AAH18088	Human cDNA sequenc
41	c 41	2.7	1902	21	AAc76936	Human cDNA sequenc
42	c 42	2.7	2252	22	AAH17566	Human chemokine al
43	c 43	2.7	2309	20	AAZ25332	Human CXC family c
44	c 44	2.7	2575	21	AAA71606	Human cDNA sequenc
45	c 45	2.7	2797	22	AAH18187	Human CXC family c
46	c 46	2.7	2799	21	AAH71607	Human cDNA sequenc
47	c 47	2.7	2845	22	AAH17249	Human cDNA sequenc
48	c 48	2.7	3585	21	AAf21279	Human low adenosin
49	c 49	2.7	3585	21	AAa35157	Human adenosine re
50	c 50	2.7	3586	22	AAH23153	Nitric oxide synth
51	c 51	2.7	3917	22	AAH18646	Human cDNA sequenc
52	c 52	2.7	9456	22	AAI63326	Human kidney relat
53	c 53	2.7	16552	22	AAI57784	Human colorectal c
54	c 54	2.7	16552	22	AAI57785	Human colorectal c
55	c 55	2.7	23241	22	AAf97870	Human neuroblastom
56	c 56	2.7	23241	22	AAf97871	Human neuroblastom
57	c 57	2.7	30417	21	AAf21282	Human low adenosin
58	c 58	2.7	30417	21	AAa35160	Human adenosine re
59	c 59	2.7	49999	20	AAZ23901	Human LOBO homolog
60	c 60	2.7	49999	20	AAZ23902	Human LOBO homolog
61	c 61	2.7	50000	21	AAA64140	Nucleotide sequenc
62	c 62	2.7	51474	22	AAf97846	Human neuroblastom
63	c 63	2.7	54548	21	AAZ45596	DNA sequence of th
64	c 64	2.7	80240	20	AAV83940	NC-contig derived
65	c 65	2.7	80595	20	AAV83939	HC-contig derived
66	c 66	2.7	84607	20	AAx90847	Human PACAP genomi
67	c 67	2.7	99960	21	AAZ50905	Human TBC-1 partia
68	c 68	2.7	121162	21	AAc66548	Human kinesin-like
69	c 69	2.7	125910	21	AAc64370	Human KCNQ5 (KCN6q
70	c 70	2.7	168575	22	AAH21613	Human hypcretin r
71	c 71	2.7	240825	22	AAf24497	Human PG-3 gene.
72	c 72	2.7	319608	21	AAH51601	Human chromosome 1
73	c 73	2.7	319608	22	AAc09301	Human schizophreni
74	c 74	2.6	131	21	AAc29646	Human secreted pro
75	c 75	2.6	182	21	AAc21820	Human secreted pro
76	c 76	2.6	182	21	AAc21823	Human secreted pro
77	c 77	2.6	183	21	AAc21824	Human secreted pro
78	c 78	2.6	215	21	AAc13839	Human secreted pro
79	c 79	2.6	278	14	AAQ61106	Human brain Expres
80	c 80	2.6	294	21	AAA52502	Human MN gene Alu-
81	c 81	2.6	294	21	AAA16583	MN genomic DNA seq
82	c 82	2.6	366	22	AAH33916	Human colon cancer
83	c 83	2.6	529	22	AAH08990	Human cDNA clone (
84	c 84	2.6	536	22	AAH09903	Human cDNA clone (

85	27	2.6	562	22	AAH10364	Human cDNA clone (158	27	2.6	141589	21	AAF21152	Human low adenosin
86	27	2.6	563	22	AAH10427	Human cDNA clone (159	27	2.6	141589	21	AAA35005	Human adenosine re
87	27	2.6	566	22	AAH10170	Human cDNA clone (160	27	2.6	141589	21	AAA35030	Human adenosine re
88	27	2.6	588	22	AAH09479	Human cDNA clone (161	27	2.6	146981	21	AAF21442	Human ELAM-1 polyn
89	27	2.6	616	20	AAH03004	Human IL-1ra BAC c	c 162	27	2.6	161425	22	AAH02340	Human AKAP10 gene
c 90	27	2.6	698	22	AAH03391	Human cDNA clone (c 163	27	2.6	162025	22	AAH02339	Human AKAP10 gene
91	27	2.6	700	22	AAH92306	Human inflammatory	164	27	2.6	209273	21	AAF21437	Human factor-relat
92	27	2.6	778	21	AAF15909	Human prostate can	165	26	2.5	176	22	AAI63308	Human kidney relat
c 93	27	2.6	1162	21	AAC74355	Human secreted pro	166	26	2.5	317	21	AAC29902	Human secreted pro
94	27	2.6	1174	21	AAA52485	Human MN gene intr	c 167	26	2.5	397	22	AAF66996	Novel human polynu
95	27	2.6	1174	21	AAA16566	Human MN intron 1	168	26	2.5	837	20	AAH37525	Human secreted pro
c 96	27	2.6	1255	21	AAC59487	Human secreted pro	c 169	26	2.5	954	21	AAC66442	Human secreted pro
c 97	27	2.6	1281	21	AAZ65092	Membrane-bound pro	c 170	26	2.5	1001	21	AAC57767	Arachidonic acid m
c 98	27	2.6	1281	21	AAZ65092	Human cDNA sequenc	c 171	26	2.5	1067	21	AAC59697	Human secreted pro
c 99	27	2.6	1281	22	AAZ65092	Human PRO1198 (UNQ	172	26	2.5	1643	22	AAH18713	Human secreted pro
c 100	27	2.6	1297	21	AAF15931	Human prostate can	c 173	26	2.5	1784	22	AAF87818	Human cDNA sequenc
c 101	27	2.6	1300	22	AAH64835	Human secreted pro	c 174	26	2.5	1968	22	AAH18031	Human haemopoietin
c 102	27	2.6	1301	22	AAH64834	Human secreted pro	c 175	26	2.5	2580	22	AAH17178	Human cDNA sequenc
c 103	27	2.6	1320	22	AAH15250	Human secreted pro	c 176	26	2.5	3292	21	AAC77356	Human ORFX ORF2911
c 104	27	2.6	1334	21	AAH81655	Human cDNA sequenc	c 177	26	2.5	3906	22	AAH73707	Human zinc finger
c 105	27	2.6	1361	22	AAF93821	N. meningitidis pa	c 178	26	2.5	15004	22	AAH27885	Nucleotide sequenc
106	27	2.6	1401	21	AAA52495	Human cDNA encodin	c 179	26	2.5	56583	21	AAF21125	Human low adenosin
107	27	2.6	1401	21	AAA16576	1.4 kb GC-rich hum	c 180	26	2.5	56583	21	AAA35003	Human adenosine re
c 108	27	2.6	1447	22	AAI59422	Human MN gene frag	c 181	26	2.5	99960	21	AAZ50905	Human TBC-1 partia
c 109	27	2.6	1548	22	AAH17022	Human polynucleoti	c 182	26	2.5	162025	22	AAH02339	Human AKAP10 gene
110	27	2.6	1609	19	AAV17088	Human cDNA sequenc	183	26	2.5	183999	22	AAF92831	Human ABC1 genomic
111	27	2.6	1613	19	AAV59193	Homo sapiens tubul	184	26	2.5	235033	19	AAV57926	Hereditary haemoch
112	27	2.6	1879	21	AAC75943	Human SR-B1 recept	185	26	2.5	237326	19	AAV57903	Hereditary haemoch
c 113	27	2.6	1901	20	AAZ01024	Human ORFX ORF1498	186	26	2.4	251	21	AAC23479	Human secreted pro
c 114	27	2.6	2084	22	AAH17430	Partial human PG1	187	25	2.4	261	21	AAC14413	Human secreted pro
c 115	27	2.6	2177	22	AAH17542	Human cDNA sequenc	188	25	2.4	307	21	AAC03264	Human secreted pro
c 116	27	2.6	2234	20	AAZ00451	Human cDNA sequenc	c 189	25	2.4	311	21	AAC25062	Human secreted pro
c 117	27	2.6	2442	22	AAH45208	Human secreted pro	c 190	25	2.4	341	22	AAF68821	Human secreted pro
c 118	27	2.6	2442	22	AAH45208	Human cytochrome C	191	25	2.4	346	21	AAC25818	Human lung tumour
c 119	27	2.6	3598	22	AAH18642	Human cDNA sequenc	192	25	2.4	827	21	AAH51124	Human secreted pro
c 120	27	2.6	4153	22	AAH18395	Human cDNA sequenc	193	25	2.4	884	21	AAH51125	Human MGST-II rela
c 121	27	2.6	4823	16	AAT03943	Human thrombopoiet	c 194	25	2.4	1289	20	AAZ40835	Human MGST-II rela
c 122	27	2.6	4823	17	AAT37384	Sequence encoding	195	25	2.4	1983	22	AAH18571	Secreted protein E
c 123	27	2.6	4823	21	AAA51993	Human thrombopoiet	c 196	25	2.4	3369	22	AAI63380	Human cDNA sequenc
c 124	27	2.6	4823	21	AAZ37785	Human thrombopoiet	c 197	25	2.4	3370	22	AAI63379	Human kidney relat
c 125	27	2.6	5117	22	AAH16663	Human cDNA sequenc	c 198	25	2.4	4696	20	AAH03041	Human kidney relat
c 126	27	2.6	6268	22	AAS01148	Interferon induced	199	25	2.4	5488	22	AAH18214	Human IL-1ra BAC c
127	27	2.6	7676	19	AAV62906	Human galactokinas	c 200	25	2.4	12312	22	AAF97872	Human cDNA sequenc
c 128	27	2.6	7849	16	AAQ94109	hML genomic DNA.	c 201	25	2.4	18647	21	AAF21059	Human neuroblastom
c 129	27	2.6	10475	9	AAN80616	Genomic clone enco	c 202	25	2.4	18648	21	AAZ34937	Human low adenosin
c 130	27	2.6	10475	9	AAN80643	Sequence of human	c 203	25	2.4	25235	22	AAH53310	Human adenosine re
131	27	2.6	10884	21	AAC66549	Human kinesin-like	c 204	25	2.4	50000	21	AAA96364	Human adenosine re
132	27	2.6	10897	17	AAT09187	MuTu putative onco	c 205	25	2.4	119950	20	AAH04994	Polymorphic repeat
133	27	2.6	10898	21	AAA52462	Human MN gene. Ho	c 206	24	2.3	129	21	AAC12085	Human yes1 gene.
134	27	2.6	10898	21	AAH16543	Human MN complete	c 207	24	2.3	129	21	AAC12085	Human secreted pro
c 135	27	2.6	10898	21	AAH16543	Human MN complete	c 208	24	2.3	210	21	AAC22572	Human secreted pro
c 136	27	2.6	11992	22	AAI63357	Human kidney relat	c 209	24	2.3	305	21	AAC02538	Human secreted pro
c 137	27	2.6	12683	22	AAI63289	Human kidney relat	210	24	2.3	411	20	AAV86226	EST clone O372. H
c 138	27	2.6	12744	22	AAI63288	Human kidney relat	211	24	2.3	466	21	AAC31674	Human secreted pro
c 139	27	2.6	14784	21	AAA64141	Nucleotide sequenc	212	24	2.3	484	22	AAH10549	Human cDNA clone (
c 140	27	2.6	21441	22	AAI62667	Human breast or ov	213	24	2.3	534	22	AAH09869	Human cDNA clone (
c 141	27	2.6	32174	22	AAI62606	Human breast or ov	c 214	24	2.3	550	22	AAH13230	Human cDNA clone (
142	27	2.6	32174	22	AAI62606	Human genomic DNA	c 215	24	2.3	777	22	AAH04994	cdNA encoding a hu
c 143	27	2.6	34488	22	AAF97854	Human neuroblastom	c 216	24	2.3	812	20	AAZ10666	Human cDNA sequenc
c 144	27	2.6	37808	20	AAH02780	Vector pMVX-BG DNA	c 217	24	2.3	1046	22	AAH18696	Human low adenosin
c 145	27	2.6	53526	19	AAT94101	Human PKD1 gene.	218	24	2.3	1520	21	AAF21107	Human adenosin
c 146	27	2.6	53577	17	AAT18551	Human polycystic k	c 219	24	2.3	1520	21	AAA34985	Human adenosine re
c 147	27	2.6	53577	19	AAT94108	Human PKD1 locus b	c 220	24	2.3	1572	21	AAC99932	Human secreted pro
148	27	2.6	56516	20	AAZ00870	PG1 genomic coding	c 221	24	2.3	1586	21	AAH31138	Human cDNA sequenc
c 149	27	2.6	56520	20	AAZ01022	Wild type PG1 codi	c 222	24	2.3	1669	22	AAH48132	Human low adenosin
c 150	27	2.6	66566	21	AAA53450	Human thioredoxin	223	24	2.3	1669	22	AAH48132	Human colon cancer
c 151	27	2.6	66685	22	AAS07380	Human genomic DNA	224	24	2.3	1827	17	AAT09865	NADH dehydrogenase
c 152	27	2.6	78925	21	AAC89888	Human FN gene. Ho	c 225	24	2.3	1827	22	AAI63631	First intron promo
153	27	2.6	106746	21	AAA10225	Human PCTA-1 genom	226	24	2.3	1961	22	AAH14525	Human kidney relat
c 154	27	2.6	121162	21	AAC66548	Human kinesin-like	c 227	24	2.3	2088	22	AAH15688	Human cDNA sequenc
155	27	2.6	122186	22	AAC89560	Human histone deac	228	24	2.3	2277	21	AAF21106	Human cDNA sequenc
156	27	2.6	138169	21	AAA34791	Human adenosine re	c 229	24	2.3	2277	21	AAH15004	Human low adenosin
157	27	2.6	141589	21	AAF20913	Human ELAM-1 polyn	c 230	24	2.3	2285	22	AAH13936	Human adenosine re
	27	2.6	141589	21	AAF21127	Human low adenosin		24	2.3	2492	22	AAH13936	Human cDNA sequenc
										2716	22	AAI63634	Human kidney relat

c 231	24	2.3	2716	22	AAI63635	Human kidney relat
232	24	2.3	2945	21	AAC68095	Human secreted pro
233	24	2.3	2945	22	AAH33162	Human colon cancer
c 234	24	2.3	3073	22	AAH15082	Human cDNA sequenc
235	24	2.3	3871	20	AAAX60801	Human secreted pro
c 236	24	2.3	4171	22	AAH16640	Human cDNA sequenc
237	24	2.3	4365	21	AAA29172	Human vanilloid re
238	24	2.3	4803	20	AAZ10062	cDNA encoding a hu
239	24	2.3	4803	20	AAZ10063	cDNA encoding a pa
240	24	2.3	4824	21	AAC66464	Human vanilloid re
c 241	24	2.3	8559	22	AAI63361	Human kidney relat
242	24	2.3	8864	21	AAF21110	Human low adenosin
243	24	2.3	8864	21	AAA34988	Human adenosine re
244	24	2.3	9734	22	AAF64136	Human lipoprotein
245	24	2.3	11617	22	AAI62937	Human genomic DNA
c 246	24	2.3	13204	20	AAZ32189	Human platelet gly
247	24	2.3	24025	17	AAT17455	Mutated BRCA1 geno
248	24	2.3	24025	17	AAT17515	Mutated BRCA1 geno
249	24	2.3	24026	17	AAT32612	BRCA1, human breas
250	24	2.3	24026	17	AAT17512	Mutated BRCA1 geno
251	24	2.3	24026	17	AAT17513	Mutated BRCA1 geno
252	24	2.3	24026	17	AAT17516	Mutated BRCA1 geno
253	24	2.3	24026	17	AAT17517	Mutated BRCA1 geno
254	24	2.3	24026	17	AAT17518	Mutated BRCA1 geno
255	24	2.3	24026	17	AAT17519	Mutated BRCA1 geno
256	24	2.3	24026	17	AAT17521	Mutated BRCA1 geno
257	24	2.3	24026	17	AAT17522	Mutated BRCA1 geno
258	24	2.3	24026	17	AAT17523	Mutated BRCA1 geno
259	24	2.3	24026	17	AAT17524	Mutated BRCA1 geno
261	24	2.3	24026	17	AAT17526	Mutated BRCA1 geno
262	24	2.3	24026	17	AAT17527	Mutated BRCA1 geno
263	24	2.3	24026	17	AAT17528	Mutated BRCA1 geno
264	24	2.3	24026	17	AAT17529	Mutated BRCA1 geno
265	24	2.3	24026	17	AAT18325	Mutated BRCA1 geno
266	24	2.3	24029	17	AAT17520	Mutated BRCA1 geno
267	24	2.3	24031	17	AAT17525	Mutated BRCA1 geno
268	24	2.3	24031	17	AAT17525	Human genomic DNA
c 269	24	2.3	32190	22	AAI62927	Human genomic DNA
c 270	24	2.3	32249	22	AAI62932	Human purH gene ge
c 271	24	2.3	41684	21	AAA28150	Human calcium sens
c 272	24	2.3	48000	22	AAF27996	Human LOBO homolog
c 273	24	2.3	49999	20	AAZ23903	Polymorphic repeat
274	24	2.3	50000	21	AAA96365	Polymorphic repeat
c 275	24	2.3	50000	21	AAA96365	Human AKAP10 gene
276	24	2.3	161425	22	AAH02340	Human secreted pro
277	23	2.2	52	21	AAC22417	Human cDNA clone (
c 278	23	2.2	351	22	AAH11840	Human brain Expres
c 279	23	2.2	374	14	AAQ59566	Human secreted pro
280	23	2.2	390	21	AAC22245	Human colon cancer
c 281	23	2.2	407	21	AAH30958	Human colon cancer
c 282	23	2.2	419	21	AAH31126	Wilson disease gen
283	23	2.2	526	16	AAQ86309	Human cDNA clone (
284	23	2.2	556	22	AAH16931	Human cDNA clone (
285	23	2.2	577	22	AAH09267	Integrin subunit b
c 286	23	2.2	688	17	AAT16856	Human colon cancer
c 287	23	2.2	698	21	AAA01622	Human inflammatory
c 288	23	2.2	700	22	AAH92897	Partial human Weel
c 289	23	2.2	1086	22	AAH45109	Human cDNA sequenc
c 290	23	2.2	1741	22	AAH15292	Human cDNA sequenc
c 291	23	2.2	1798	22	AAH13835	Human secreted pro
c 292	23	2.2	1871	22	AAH64770	Human cDNA sequenc
c 293	23	2.2	2504	22	AAH17838	Novel protein kina
c 294	23	2.2	2615	22	AAF44670	Human NIP2b cDNA.
c 295	23	2.2	3065	21	AAD00128	Human cDNA sequenc
c 296	23	2.2	3227	22	AAH16625	Human ORFX ORF1316
297	23	2.2	5124	21	AAC75761	Human neuroblastom
c 298	23	2.2	6405	22	AAF97850	Human KLK-L4 gene.
299	23	2.2	10080	21	AAA95933	Genomic sequence e
c 300	23	2.2	11715	9	AAN81564	Complete genomic s
c 301	23	2.2	11724	8	AAN70102	Human liver cell s
c 302	23	2.2	11724	22	AAH57497	Human protein C ge
c 303	23	2.2	11725	17	AAT32796	
11725	18	2.2	11725	20	AAZ79723	Human protein C ge
11725	20	2.2	11725	20	AAZ32179	Human protein C nu
11725	22	2.2	11725	22	AAF54051	Human death-associ
36221	22	2.2	36221	22	AA500624	Human purH gene ge
41684	21	2.2	41684	21	AAA28150	Nucleotide sequenc
44848	21	2.2	44848	21	AAZ75080	Olfactory receptor
144460	21	2.2	144460	21	AAZ93815	Human ABC1 genomic
183999	22	2.2	183999	22	AAF92831	Human secreted pro
117	21	2.1	117	21	AAC12107	Human secreted pro
125	21	2.1	125	21	AAC14035	Human secreted pro
183	21	2.1	183	21	AAC20144	Human secreted pro
189	16	2.1	189	16	AAT24873	Human gene signatu
249	21	2.1	249	21	AAC21715	Human secreted pro
289	21	2.1	289	21	AAC04999	Human secreted pro
293	21	2.1	293	21	AAC26337	Human secreted pro
300	20	2.1	300	20	AAZ13542	Human gene express
302	21	2.1	302	21	AAC02929	Human secreted pro
334	21	2.1	334	21	AAC19789	Human secreted pro
357	21	2.1	357	21	AAC16430	Human secreted pro
365	21	2.1	365	21	AAC03488	Human secreted pro
365	21	2.1	365	21	AAC15064	Expressed Sequence
382	14	2.1	382	14	AAQ39885	Human brain Expres
382	14	2.1	382	14	AAQ59297	Human secreted pro
495	22	2.1	495	22	AAD08525	Human secreted exp
527	21	2.1	527	21	AAA45379	N. meningitidis pa
540	21	2.1	540	21	AAA81688	Human cDNA clone (
559	22	2.1	559	22	AAH10229	Human cDNA clone (
570	22	2.1	570	22	AAH13418	Human inflammatory
700	22	2.1	700	22	AAH92231	Human inflammatory
778	22	2.1	778	22	AAH07224	Human cDNA clone (
789	21	2.1	789	21	AAC93467	Human secreted pro
843	22	2.1	843	22	AAH04879	Human cDNA clone (
877	19	2.1	877	19	AAV04635	Human cytoplasmic
1001	21	2.1	1001	21	AAC57479	Arachidonic acid m
1098	19	2.1	1098	19	AAV24143	Homo sapiens BARD1
1495	22	2.1	1495	22	AAD08502	Human secreted pro
1533	21	2.1	1533	21	AAF21366	Human low adenosin
1533	21	2.1	1533	21	AAA35244	Human adenosine re
1774	22	2.1	1774	22	AAH99647	Human protein enco
1784	21	2.1	1784	21	AAZ36778	Nucleotide sequenc
1816	22	2.1	1816	22	AAF72758	Human prostate can
1848	22	2.1	1848	22	AAH16421	Human cDNA sequenc
1880	22	2.1	1880	22	AAH16035	Human cDNA sequenc
1964	22	2.1	1964	22	AAH15328	Human cDNA sequenc
2157	21	2.1	2157	21	AAZ74405	Human secreted pro
2176	22	2.1	2176	22	AAH15589	Human cDNA sequenc
2243	22	2.1	2243	22	AAC89709	Human oligodendroc
2271	22	2.1	2271	22	AAH17984	Human cDNA sequenc
2325	22	2.1	2325	22	AAD07578	Human secreted pro
2337	21	2.1	2337	21	AAZ39581	Human cREL nucleot
2350	22	2.1	2350	22	AAH15637	Human cDNA sequenc
2350	22	2.1	2350	22	AAH15637	Human polynucleoti
2548	22	2.1	2548	22	AAH17488	Human cDNA sequenc
2633	22	2.1	2633	22	AAH17293	Human cDNA sequenc
3161	22	2.1	3161	22	AAH21368	Human low adenosin
3191	21	2.1	3191	21	AAF35246	Human adenosine re
3303	22	2.1	3303	22	AAH18287	Human cDNA sequenc
3885	19	2.1	3885	19	AAV25508	Human bax promoter
4041	22	2.1	4041	22	AAH46858	Human serine/threo
4156	22	2.1	4156	22	AAH57458	Human liver cell s
4157	22	2.1	4157	22	AAH57457	Human liver cell s
6775	22	2.1	6775	22	AAI63671	Human kidney relat
7620	14	2.1	7620	14	AAQ39286	Glucocerebrosidase
11992	22	2.1	11992	22	AAI63357	Human kidney relat
12151	21	2.1	12151	21	AAZ36757	DNA sequence encod
14417	22	2.1	14417	22	AAI62923	Human genomic DNA
14426	22	2.1	14426	22	AAI62921	Human genomic DNA
14426	22	2.1	14426	22	AAI62924	Human genomic DNA
14448	22	2.1	14448	22	AAI62922	Human genomic DNA
14451	22	2.1	14451	22	AAI62925	Human HCR DNA. HO
14483	22	2.1	14483	22	AAH45311	Survivin gene. HO
14796	19	2.1	14796	19	AAV27941	
304	23	2.2	304	23		
305	23	2.2	305	23		
306	23	2.2	306	23		
307	23	2.2	307	23		
308	23	2.2	308	23		
309	23	2.2	309	23		
310	23	2.2	310	23		
311	23	2.2	311	23		
312	22	2.1	312	22		
313	22	2.1	313	22		
314	22	2.1	314	22		
315	22	2.1	315	22		
316	22	2.1	316	22		
317	22	2.1	317	22		
318	22	2.1	318	22		
319	22	2.1	319	22		
320	22	2.1	320	22		
321	22	2.1	321	22		
322	22	2.1	322	22		
323	22	2.1	323	22		
324	22	2.1	324	22		
325	22	2.1	325	22		
326	22	2.1	326	22		
327	22	2.1	327	22		
328	22	2.1	328	22		
329	22	2.1	329	22		
330	22	2.1	330	22		
331	22	2.1	331	22		
332	22	2.1	332	22		
333	22	2.1	333	22		
334	22	2.1	334	22		
335	22	2.1	335	22		
336	22	2.1	336	22		
337	22	2.1	337	22		
338	22	2.1	338	22		
339	22	2.1	339	22		
340	22	2.1	340	22		
341	22	2.1	341	22		
342	22	2.1	342	22		
343	22	2.1	343	22		
344	22	2.1	344	22		
345	22	2.1	345	22		
346	22	2.1	346	22		
347	22	2.1	347	22		
348	22	2.1	348	22		
349	22	2.1	349	22		
350	22	2.1	350	22		
351	22	2.1	351	22		
352	22	2.1	352	22		
353	22	2.1	353	22		
354	22	2.1	354	22		
355	22	2.1	355	22		
356	22	2.1	356	22		
357	22	2.1	357	22		
358	22	2.1	358	22		
359	22	2.1	359	22		
360	22	2.1	360	22		
361	22	2.1	361	22		

523	20	1.9	603	22	AAH03596	Human cDNA clone (
524	20	1.9	614	20	AAV88921	EST clone HW591.
525	20	1.9	657	22	AAD07935	Human TDC protein-
526	20	1.9	674	21	AAC59543	Human secreted pro
527	20	1.9	676	21	AAH72487	Human pancreatic d
528	20	1.9	691	22	AAH03326	Human cDNA clone (
529	20	1.9	700	22	AAH92015	Human inflammatory
530	20	1.9	700	22	AAH92368	Human inflammatory
531	20	1.9	700	22	AAH92369	Human inflammatory
532	20	1.9	700	22	AAH92682	Human inflammatory
533	20	1.9	700	22	AAH07356	Human cDNA clone (
534	20	1.9	705	22	AAH03213	Human cDNA clone (
535	20	1.9	719	20	AAV88412	EST clone EK491.
536	20	1.9	740	20	AAZ15062	Human gene express
537	20	1.9	748	20	AAZ24829	Human secreted pro
538	20	1.9	748	22	AAH08166	Human cDNA clone (
539	20	1.9	779	21	AAC59174	Human secreted pro
540	20	1.9	796	20	AAH37484	Human secreted pro
541	20	1.9	799	22	AAH04675	Human cDNA clone (
542	20	1.9	804	22	AAH04563	Human cDNA clone (
543	20	1.9	804	22	AAH06560	Human cDNA clone (
544	20	1.9	811	22	AAH73279	Human cervical can
545	20	1.9	847	22	AAD08516	Human secreted pro
546	20	1.9	847	22	AAD05451	Human secreted pro
547	20	1.9	919	20	AAH61340	DNA encoding a hum
548	20	1.9	942	20	AAH81838	Human malignancy-a
549	20	1.9	960	21	AAC79789	Human secreted pro
550	20	1.9	968	22	AAH98581	Human EST-derived
551	20	1.9	976	22	AAI63545	Human kidney relat
552	20	1.9	992	21	AAC59452	Human secreted pro
553	20	1.9	1001	21	AAC57825	Arachidonic acid m
554	20	1.9	1009	22	AAI57640	Human colorectal c
555	20	1.9	1017	20	AAH97756	Extended human sec
556	20	1.9	1028	21	AAH21761	Human breast and o
557	20	1.9	1086	21	AAC74264	Human secreted pro
558	20	1.9	1138	21	AAC79743	Human secreted pro
559	20	1.9	1160	20	AAH84540	Human secreted pro
560	20	1.9	1162	21	AAC59157	Human foetal cDNA,
561	20	1.9	1222	22	AAH94385	Human secreted pro
562	20	1.9	1227	22	AAH45078	Human secreted pro
563	20	1.9	1251	21	AAC59177	Human secreted pro
564	20	1.9	1252	21	AAH22419	Human secreted pro
565	20	1.9	1320	21	AAH22398	Human secreted pro
566	20	1.9	1321	22	AAH19195	Human secreted pro
567	20	1.9	1331	22	AAH59595	Human cell cycle a
568	20	1.9	1334	22	AAH91877	Human secreted pro
569	20	1.9	1380	22	AAH44674	Novel protein kina
570	20	1.9	1413	21	AAD00684	Human Hydrolase pr
571	20	1.9	1419	22	AAI60860	Human polynucleoti
572	20	1.9	1425	22	AAH64215	Human secreted pro
573	20	1.9	1426	22	AAI59074	Human polynucleoti
574	20	1.9	1465	22	AAH17214	Human cDNA sequenc
575	20	1.9	1470	14	AAQ47355	Myotonic dystrophy
576	20	1.9	1559	16	AAH02714	MART-1 melanoma an
577	20	1.9	1559	20	AAZ07987	Human melanoma ant
578	20	1.9	1559	20	AAZ20065	Human MART1 melano
579	20	1.9	1561	22	AAH13688	Human cDNA sequenc
580	20	1.9	1579	22	AAH17258	Human cDNA sequenc
581	20	1.9	1617	22	AAH15286	Human secreted pro
582	20	1.9	1643	21	AAH87754	cDNA encoding huma
583	20	1.9	1643	22	AAH64036	Human secreted pro
584	20	1.9	1657	22	AAH19175	Human secreted pro
585	20	1.9	1679	22	AAH17499	Human cDNA sequenc
586	20	1.9	1687	15	AAQ54642	Human A2b adenosin
587	20	1.9	1687	21	AAH20857	Human adenosine A2
588	20	1.9	1687	21	AAH20868	Human adenosine re
589	20	1.9	1687	21	AAA34746	Human adenosine re
590	20	1.9	1687	21	AAH14703	Human cDNA sequenc
591	20	1.9	1696	22	AAH14703	Human cDNA sequenc
592	20	1.9	1737	22	AAH94472	Human hydrophobic
593	20	1.9	1756	22	AAH15180	Human cDNA sequenc
594	20	1.9	1762	22	AAD08495	Human secreted pro
595	20	1.9	1772	22	AAH18437	Human cDNA sequenc
1811	22	1.9	1811	22	AAD05407	Human secreted pro
1821	22	1.9	1821	22	AAH18062	Human cDNA sequenc
1873	21	1.9	1873	21	AAA39056	Human secreted pro
1897	21	1.9	1897	21	AAA39083	Human secreted pro
1916	22	1.9	1916	22	AAH44418	Human dihydroorota
1949	22	1.9	1949	22	AAH17260	Human cDNA sequenc
1954	22	1.9	1954	22	AAH17532	Human cDNA sequenc
1984	22	1.9	1984	22	AAH16531	Human cDNA sequenc
2005	22	1.9	2005	22	AAH13890	Sequence flanking
2040	17	1.9	2040	17	AAH43949	Fucosyltransferase
2042	18	1.9	2042	18	AAH80111	Glycosyltransferas
2043	12	1.9	2043	12	AAQ13330	DNA encoding a gly
2043	15	1.9	2043	15	AAQ56906	DNA encoding a gly
2043	18	1.9	2043	18	AAH76769	Human alpha 1,3/4
2043	18	1.9	2043	18	AAH16765	Human alpha(1,3/1,
2065	22	1.9	2065	22	AAD05414	Human secreted pro
2070	22	1.9	2070	22	AAI59065	Human polynucleoti
2116	22	1.9	2116	22	AAH14779	Human cDNA sequenc
2135	22	1.9	2135	22	AAH13650	Human cDNA sequenc
2148	22	1.9	2148	22	AAH15766	Human cDNA sequenc
2152	22	1.9	2152	22	AAH18693	Polynucleotide seq
2156	20	1.9	2156	20	AAZ10807	Human cDNA sequenc
2195	22	1.9	2195	22	AAH18440	DNA ligase 10 cDNA
2209	22	1.9	2209	22	AAH46146	Human cDNA sequenc
2246	22	1.9	2246	22	AAH16261	Human cDNA sequenc
2266	22	1.9	2266	22	AAH18640	Human polynucleoti
2296	22	1.9	2296	22	AAI60086	Human secreted pro
2316	22	1.9	2316	22	AAD05468	DNA encoding human
2326	20	1.9	2326	20	AAH85833	Human mammatatin
2326	22	1.9	2326	22	AAH31279	Human secreted pro
2351	22	1.9	2351	22	AAH17651	Human cDNA sequenc
2355	22	1.9	2355	22	AAH31280	Human mammatatin
2380	19	1.9	2380	19	AAV34456	Human MHC class I
2400	22	1.9	2400	22	AAH15537	Human cDNA sequenc
2417	19	1.9	2417	19	AAV24305	Human mammary cell
2418	22	1.9	2418	22	AAH17392	Human cDNA sequenc
2418	22	1.9	2418	22	AAH31278	Human mammatatin
2498	22	1.9	2498	22	AAH18310	Human cDNA sequenc
2739	22	1.9	2739	22	AAH17762	Human cDNA sequenc
2777	19	1.9	2777	19	AAV69329	Human LIR-18A3 CDN
2777	21	1.9	2777	21	AAA54592	Leukocyte immunogl
2790	19	1.9	2790	19	AAV38608	Human monocyte gen
2823	22	1.9	2823	22	AAH16591	Human cDNA sequenc
2913	22	1.9	2913	22	AAH14622	Human cDNA sequenc
2922	19	1.9	2922	19	AAH69328	Human LIR-P3G2 CDN
2922	21	1.9	2922	21	AAA54591	Leukocyte immunogl
2950	15	1.9	2950	15	AAQ58744	AMEF2. Homo sapien
2968	15	1.9	2968	15	AAQ58740	MEF2. Homo sapien
3051	22	1.9	3051	22	AAH17317	Human cDNA sequenc
3137	22	1.9	3137	22	AAH17117	Human cDNA sequenc
3170	22	1.9	3170	22	AAH17301	Human cDNA sequenc
3198	22	1.9	3198	22	AAH17881	Human cDNA sequenc
3240	22	1.9	3240	22	AAH64150	Part of human lipo
3332	21	1.9	3332	21	AAZ88392	Human ZAP-3 encodi
3337	21	1.9	3337	21	AAC99857	Human secreted pro
3345	20	1.9	3345	20	AAH78421	Human HPT-1 protei
3345	22	1.9	3345	22	AAI29511	Human intestinal p
3415	21	1.9	3415	21	AAC59802	Human secreted pro
3568	20	1.9	3568	20	AAV99924	DY11 torsion dysto
3597	21	1.9	3597	21	AAH20377	Human low adenosin
3597	21	1.9	3597	21	AAA34855	Human adenosine re
3612	22	1.9	3612	22	AAH98579	Human EST-derived
3612	22	1.9	3612	22	AAI57908	Human polynucleoti
3621	19	1.9	3621	19	AAV33072	Human polynucleoti
3666	22	1.9	3666	22	AAH18651	Dendritic cell-der
4035	22	1.9	4035	22	AAH16648	Human cDNA sequenc
4459	20	1.9	4459	20	AAV99094	Human cDNA sequenc
4792	22	1.9	4792	22	AAI63544	DNA methyltransfer
4941	22	1.9	4941	22	AAI63251	Human kidney relat
5243	22	1.9	5243	22	AAH18240	Human cDNA sequenc

c 669	20	1.9	5446	22	AAI63390	Human kidney relat
c 670	20	1.9	5830	22	AAI63546	Human kidney relat
c 671	20	1.9	6045	22	AAH18244	Human cDNA sequenc
c 672	20	1.9	6210	15	AAQ63815	c-fos gene. Homo
c 673	20	1.9	6210	21	AAF21091	Human low adenodin
c 674	20	1.9	6210	21	AAA34969	Human adenosine re
c 675	20	1.9	6210	22	AAH18663	Human cDNA sequenc
c 676	20	1.9	6235	21	AAZ29169	Human G-CSF genom
c 677	20	1.9	6312	22	AAI62822	Partial sequence o
c 678	20	1.9	6679	21	AAZ29170	Human low adenodin
c 679	20	1.9	6711	21	AAF21122	Human adenosine re
c 680	20	1.9	6711	21	AAA35000	Human polynucleoti
c 681	20	1.9	6799	22	AAI58419	Human adenosine A2
c 682	20	1.9	7144	21	AAF21439	Insulin receptor g
c 683	20	1.9	7240	22	AAH50570	Human insulin rece
c 684	20	1.9	7240	22	AAH31170	Human insulin rece
c 685	20	1.9	7240	22	AAH31267	Human insulin rece
c 686	20	1.9	7240	22	AAD04467	Human tumour suppr
c 687	20	1.9	7397	19	AAV60578	Human genomic DNA
c 688	20	1.9	7733	22	AAI62831	Methods for diagno
c 689	20	1.9	8205	22	AAI62948	Human low adenodin
c 690	20	1.9	8353	18	AAT89396	Human adenosine re
c 691	20	1.9	9203	21	AAF21350	Human kidney relat
c 692	20	1.9	9203	21	AAA35228	Human neuroblastom
c 693	20	1.9	9439	22	AAI63723	Human low adenodin
c 694	20	1.9	9799	22	AAF97860	Human adenosine re
c 695	20	1.9	10032	21	AAI63723	Human neuroblastom
c 696	20	1.9	10032	21	AAA35002	Human low adenodin
c 697	20	1.9	10120	20	AAZ23683	Human adenosine re
c 698	20	1.9	10380	18	AAT67164	Human DK1 DNA fra
c 699	20	1.9	10825	18	AAT96690	Human alpha-N-acet
c 700	20	1.9	10825	22	AAC68425	Hereditary haemoch
c 701	20	1.9	10825	22	AAC68426	Human hereditary h
c 702	20	1.9	10825	22	AAC68427	Human hereditary h
c 703	20	1.9	10825	22	AAC68428	Human hereditary h
c 704	20	1.9	11531	15	AAQ54222	BSS1/CEL Gene. Ho
c 705	20	1.9	11531	20	AZ20299	Human bile salt-st
c 706	20	1.9	11570	21	AAA95905	Human KLK-L2 gene.
c 707	20	1.9	12146	21	AAA96794	Genomic DNA of a h
c 708	20	1.9	13747	14	AAQ47342	Myotonic dystrophy
c 709	20	1.9	14919	22	AAI63992	Human polynucleoti
c 710	20	1.9	15848	20	AZ32190	Human heparin cofa
c 711	20	1.9	16225	22	AAI62650	Human breast or ov
c 712	20	1.9	16595	21	AAF21095	Human low adenodin
c 713	20	1.9	16595	21	AAA34973	Human adenosine re
c 714	20	1.9	17844	21	AAI21358	Human low adenodin
c 715	20	1.9	17844	21	AAA35236	Human adenosine re
c 716	20	1.9	19440	20	AAV99129	DNA methyltransfer
c 717	20	1.9	20268	22	AAI57636	Human colorectal c
c 718	20	1.9	22081	22	AAF97862	Human neuroblastom
c 719	20	1.9	23603	22	AAI62936	Human genomic DNA
c 720	20	1.9	23613	22	AAI62930	Human genomic DNA
c 721	20	1.9	24025	17	AAT17455	Mutated BRCA1 geno
c 722	20	1.9	24025	17	AAT17515	Mutated BRCA1 geno
c 723	20	1.9	24026	17	AAT32612	BRCA1, human breas
c 724	20	1.9	24026	17	AAT17512	Mutated BRCA1 geno
c 725	20	1.9	24026	17	AAT17513	Mutated BRCA1 geno
c 726	20	1.9	24026	17	AAT17514	Mutated BRCA1 geno
c 727	20	1.9	24026	17	AAT17516	Mutated BRCA1 geno
c 728	20	1.9	24026	17	AAT17517	Mutated BRCA1 geno
c 729	20	1.9	24026	17	AAT17518	Mutated BRCA1 geno
c 730	20	1.9	24026	17	AAT17519	Mutated BRCA1 geno
c 731	20	1.9	24026	17	AAT17521	Mutated BRCA1 geno
c 732	20	1.9	24026	17	AAT17522	Mutated BRCA1 geno
c 733	20	1.9	24026	17	AAT17523	Mutated BRCA1 geno
c 734	20	1.9	24026	17	AAT17524	Mutated BRCA1 geno
c 735	20	1.9	24026	17	AAT17526	Mutated BRCA1 geno
c 736	20	1.9	24026	17	AAT17527	Mutated BRCA1 geno
c 737	20	1.9	24026	17	AAT17528	Mutated BRCA1 geno
c 738	20	1.9	24026	17	AAT17529	Mutated BRCA1 geno
c 739	20	1.9	24026	17	AAT17530	BRCA1 genomic sequ
c 740	20	1.9	24026	17	AAT18325	BRCA1, human breas
c 741	20	1.9	24029	17	AAT17520	Mutated BRCA1 geno
c 742	20	1.9	24031	17	AAT17525	Mutated BRCA1 geno
c 743	20	1.9	24218	22	AAI62935	Human genomic DNA
c 744	20	1.9	26006	22	AAI62592	Human breast or ov
c 745	20	1.9	26006	22	AAI63583	Human kidney relat
c 746	20	1.9	26013	22	AAI62590	Human breast or ov
c 747	20	1.9	26013	22	AAI63581	Human kidney relat
c 748	20	1.9	26018	22	AAI62591	Human breast or ov
c 749	20	1.9	26018	22	AAI63582	Human kidney relat
c 750	20	1.9	26928	20	AAZ32184	Human prothrombin
c 751	20	1.9	32152	22	AAI57791	Human colorectal c
c 752	20	1.9	32183	22	AAI63617	Human colorectal c
c 753	20	1.9	32204	22	AAI57790	Human kidney relat
c 754	20	1.9	33780	22	AAH24652	Nucleotide sequenc
c 755	20	1.9	34488	22	AAF97854	Human neuroblastom
c 756	20	1.9	38186	20	AAZ32028	Human METH1 relate
c 757	20	1.9	38186	22	AAC90085	AC004449 cDNA clon
c 758	20	1.9	39198	22	AAF58067	Human polyamine mo
c 759	20	1.9	44453	20	AAZ23519	Human kidney amino
c 760	20	1.9	49999	20	AAZ23902	Human LOBO homolog
c 761	20	1.9	49999	20	AAZ23904	Human LOBO homolog
c 762	20	1.9	50000	21	AAA96363	Polymorphic repeat
c 763	20	1.9	50000	21	AAA64139	Nucleotide sequenc
c 764	20	1.9	50000	22	AAF54867	Nucleotide sequenc
c 765	20	1.9	51474	22	AAF97846	Human neuroblastom
c 766	20	1.9	52216	22	AAH28355	Nucleotide sequenc
c 767	20	1.9	72928	20	AAZ18355	Human neuroblastom
c 768	20	1.9	72928	21	AAH80253	Nucleotide sequenc
c 769	20	1.9	80331	22	AAC89559	Human ASTH1J 5' ge
c 770	20	1.9	80331	22	AAC89559	Human histone deac
c 771	20	1.9	81001	22	AAF30035	Human histone deac
c 772	20	1.9	81145	22	AAF54868	Human apolipoprote
c 773	20	1.9	84607	20	AAZ90847	Genomic nucleotide
c 774	20	1.9	90050	21	AAZ91925	Human PACAP genomi
c 775	20	1.9	97662	22	AAF83908	Wild type (c57BL/6
c 776	20	1.9	110000	22	AAF83908	Genomic sequence o
c 777	20	1.9	114793	22	AAD08215	Nucleotide sequenc
c 778	20	1.9	114793	22	AAD08215	Human genome from
c 779	20	1.9	117609	21	AAF21435	Human genome from
c 780	20	1.9	162450	21	AAZ86967	Human receptor-rel
c 781	20	1.9	235033	19	AAV57926	Retinoblastoma bin
						Hereditary haemoch

ALIGNMENTS

RESULT	1
AAD03816	
ID	AAD03816 standard; cDNA; 1041 BP.
XX	
AC	AAD03816;
XX	
DT	19-JUN-2001 (first entry)
XX	
DE	Human kinase cDNA #5.
XX	
KW	Human; kinase; gene therapy; bioreactor; mental disorder;
KW	biological disorder; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	1..1041
FT	/*tag= a
FT	/product= "Human kinase #5"
FT	/note= "The coding region does not include stop codon"
XX	
XX	
PN	WO200123579-A1.
XX	
PD	05-APR-2001.
XX	
PF	27-SEP-2000; 2000WO-US26621.
XX	


```
XX 28-SEP-1999; 99US-0156511.
XX (LEXI-) LEXICON GENETICS INC.
XX Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
XX WPI; 2001-266166/27.
DR P-PSDB; AAE00491.
XX
XX New isolated human kinase polynucleotide useful for generating
XX antibodies, as reagents in diagnostic assays and for screening for
XX compounds useful for treating mental, biological or medical diseases
XX
XX Disclosure; Page 28; 38pp; English.
XX
XX The present sequence is a cDNA encoding novel human protein
XX (NHP) known as human kinase. The human kinases share structural
XX similarity with animal kinases, more particularly serine or
XX threonine protein kinases. Human kinase cDNA is useful for the
XX detection of mutant human kinase for the diagnosis of disease,
XX and also as a therapeutic. It is useful for screening drugs
XX effective in the treatment of symptomatic or phenotypic
XX manifestations perturbing the normal function of NHP in the
XX body. The NHP nucleotide sequences are useful for generation of
XX antibodies, as reagents in diagnostic assays, for the
XX identification of other cellular gene products related to human
XX kinases, and as reagents in assays for screening compounds that
XX are useful for treating mental, biological or medical disorders.
XX NHP oligonucleotides are used as probes. The labelled NHP probes
XX are useful for screening human genomic library for identifying
XX polymorphisms and as primers in amplification assays to detect
XX mutations within the exons, introns and splice sites that can
XX be used in diagnostics and pharmacogenomics. Nucleotide construct
XX encoding NHP products are used to genetically engineer cells
XX in vivo that functions as bioreactors in the body delivering a
XX continuous supply of NHP to the body. Nucleotide constructs
XX encoding functional NHPs are used in gene therapy for the
XX modulation of NHP expression.
XX
XX Sequence 1068 BP; 344 A; 197 C; 240 G; 287 T; 0 other;
```

```
Query Match      98.5%; Score 1025; DB 22; Length 1068;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 atggaaaagtatgaaaaattagctaagactggagaagggtcttatgggtgtattcaaa 60
Db 1 atggaaaagtatgaaaaattagctaagactggagaagggtcttatgggtgtattcaaa 60
QY 61 tgcagaaacaaaacctctggacaagtagtagctgttataaaaatttgggaatctgaagat 120
Db 61 tgcagaaacaaaacctctggacaagtagtagctgttataaaaatttgggaatctgaagat 120
QY 121 gattcctgttgaagaaaatagcactaagagaaatacgtatgttgaagcaattataaacat 180
Db 121 gattcctgttgaagaaaatagcactaagagaaatacgtatgttgaagcaattataaacat 180
QY 181 ccaaatcttgtgaacctcatcgagggtgttcaggagaaaaaggaaaaatgcatttagtttt 240
Db 181 ccaaatcttgtgaacctcatcgagggtgttcaggagaaaaaggaaaaatgcatttagtttt 240
QY 241 gaatactgtgatcatcacacttttaaatgagctggagaaaccccaatggagttgctgat 300
Db 241 gaatactgtgatcatcacacttttaaatgagctggagaaaccccaatggagttgctgat 300
QY 301 ggagtgatcaaaagcgtattatggcaaacacttcaagctcttaatttctgtcatatacat 360
Db 301 ggagtgatcaaaagcgtattatggcaaacacttcaagctcttaatttctgtcatatacat 360
QY 361 aactgtattcacagagatatataaacctgaaaaatttctaataactaagcaaggaataatc 420
Db 361 aactgtattcacagagatatataaacctgaaaaatttctaataactaagcaaggaataatc 420
```

```
Db 361 aactgtattcacagagatatataaacctgaaaaatttctaataactaagcaaggaataatc 420
QY 421 aagatttgtgacttcgggtttgcacaaaattctgattccaggagatgcctacaccgattat 480
Db 421 aagatttgtgacttcgggtttgcacaaaattctgattccaggagatgcctacaccgattat 480
QY 481 gtagctacgagatggtaccgagctcctgaacttcttgtgggagatactcagtatggttct 540
Db 481 gtagctacgagatggtaccgagctcctgaacttcttgtgggagatactcagtatggttct 540
QY 541 tcagtcgatatatgggctattggttggtttttgcagagctcctgacagggccagccactg 600
Db 541 tcagtcgatatatgggctattggttggtttttgcagagctcctgacagggccagccactg 600
QY 601 tggcctggaaaaatcagatgtggaccaactttatctgataatcagaacactagtagagacg 660
Db 601 tggcctggaaaaatcagatgtggaccaactttatctgataatcagaacactagtagagacg 660
QY 661 gggtttcgccatgttgaccaggctgtgtctcgaactcttgacgtcaagtgatccacctgcc 720
Db 661 gggtttcgccatgttgaccaggctgtgtctcgaactcttgacgtcaagtgatccacctgcc 720
QY 721 gtagcctctcaaaagtctggaattacaggaaaaattaatcccaagacatcaatcaatcttt 780
Db 721 gtagcctctcaaaagtctggaattacaggaaaaattaatcccaagacatcaatcaatcttt 780
QY 781 aaaagtaaagggtttttccatggcatcagtatcacctgagccagaaacatggaactctt 840
Db 781 aaaagtaaagggtttttccatggcatcagtatcacctgagccagaaacatggaactctt 840
QY 841 gaggaaaagttctcagatgttccatcctctgtggctctgaacttcattcatggtgtctgaag 900
Db 841 gaggaaaagttctcagatgttccatcctctgtggctctgaacttcattcatggtgtctgaag 900
QY 901 atgaatccagatgacagattaacctgttcccaactcctggagagctcctacttattct 960
Db 901 atgaatccagatgacagattaacctgttcccaactcctggagagctcctacttattct 960
QY 961 ttccaagaggcccaaaattaaaagaaaaagcacgtaataaggaaagaaacagagacgcca 1020
Db 961 ttccaagaggcccaaaattaaaagaaaaagcacgtaataaggaaagaaacagagacgcca 1020
QY 1021 caggt 1025
Db 1021 caggt 1025
```

```
RESULT 3
AAD03817
ID AAD03817 standard; cDNA; 945 BP.
XX
AC AAD03817;
XX
DT 19-JUN-2001 (first entry)
XX
DE Human kinase cDNA #6.
XX
KW Human; kinase; gene therapy; bioreactor; mental disorder;
XX biological disorder; ss.
XX
OS Homo sapiens.
XX
FH Key
FT CDS
FT 1..945
FT /*tag= a
FT /product= "Human kinase #6"
FT /note= "The coding region does not include stop codon"
XX
PN WO200123579-A1.
XX
PD 05-APR-2001.
XX
```


CC identification of other cellular gene products related to human
CC kinases, and as reagents in assays for screening compounds that
CC are useful for treating mental, biological or medical disorders.
CC NHP oligonucleotides are used as probes. The labelled NHP probes
CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.
XX
SQ Sequence 972 BP; 323 A; 172 C; 214 G; 263 T; 0 other;

Query Match 62.6%; Score 652; DB 22; Length 972;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 652; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atggaaaagtataaaaaattagctaagactggagaagggtcttatggggtgtattcaaa 60
Db |
1 atggaaaagtataaaaaattagctaagactggagaagggtcttatggggtgtattcaaa 60
QY 61 tgcagaaaacaaaacctctgtgacaaagttagctgtataaaaaatttgtggaatctgaagat 120
Db |
61 tgcagaaaacaaaacctctgtgacaaagttagctgtataaaaaatttgtggaatctgaagat 120
QY 121 gatcctgtgttaagaaaatagcactaagaagatacgtatgttgaagcaaatataaacat 180
Db |
121 gatcctgtgttaagaaaatagcactaagaagatacgtatgttgaagcaaatataaacat 180
QY 181 ccaaatcttctgaacctcatcaggtgttcaggagaaaaagcaaaatgcatttagtttt 240
Db |
181 ccaaatcttctgaacctcatcaggtgttcaggagaaaaagcaaaatgcatttagtttt 240
QY 241 gaatactgtgatcatcacacttttaaatgagctggaagaaaccccaaatggagttgctgat 300
Db |
241 gaatactgtgatcatcacacttttaaatgagctggaagaaaccccaaatggagttgctgat 300
QY 301 ggagtgatcaaaaagcgtattatggcaaacacttcaagctcttaatttctgtcatatacat 360
Db |
301 ggagtgatcaaaaagcgtattatggcaaacacttcaagctcttaatttctgtcatatacat 360
QY 361 aactgtattcacagagatatataaacctgaaaatatttctaataactaagcaaggaaataatc 420
Db |
361 aactgtattcacagagatatataaacctgaaaatatttctaataactaagcaaggaaataatc 420
QY 421 aagatttgtgacttcgggtttgcacaaaattctgattccaggagagatgcctacacggattat 480
Db |
421 aagatttgtgacttcgggtttgcacaaaattctgattccaggagagatgcctacacggattat 480
QY 481 gtagctacgagatggtaccgagctcctgaacttcttgtgggagatactcagtatggttct 540
Db |
481 gtagctacgagatggtaccgagctcctgaacttcttgtgggagatactcagtatggttct 540
QY 541 tcagtcgatataatgggctattggtgtgttttgcagagctcctgcacaggccagccactg 600
Db |
541 tcagtcgatataatgggctattggtgtgttttgcagagctcctgcacaggccagccactg 600
QY 601 tggcctggaaaatcagatgtggaccactttatctgataatcagaacactag 652
Db |
601 tggcctggaaaatcagatgtggaccactttatctgataatcagaacactag 652

RESULT 5
AAD03818
ID AAD03818 standard; cDNA; 1819 BP.
XX
AC AAD03818;
XX
DT 19-JUN-2001 (first entry)

XX Human kinase cDNA #7.
DE
XX
KW Human; kinase; gene therapy; bioreactor; mental disorder;
KW biological disorder; polymorphism; ss.
XX Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT 5'UTR 1..137
FT /tag= a
FT CDS 138..701
FT /tag= b
FT /product= "Human Kinase #1"
FT 3'UTR 702..1819
FT /tag= c
FT misc_difference 1684
FT /tag= d
FT /note= "polymorphism site; given as r in the sequence"
XX
PN WO200123579-A1.
XX
PD 05-APR-2001.
XX
PF 27-SEP-2000; 2000WO-US26621.
XX
PR 28-SEP-1999; 99US-0156511.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
XX
DR WPI; 2001-266166/27.
DR P-PSDB; AAE00490.
XX
PT New isolated human kinase polynucleotide useful for generating
PT antibodies, as reagents in diagnostic assays and for screening for
PT compounds useful for treating mental, biological or medical diseases -
PS Disclosure; Page 34-35; 38pp; English.
XX
CC The present sequence encodes a novel human protein (NHP) which
CC is a human kinase. A polymorphism was identified in the 3' UTR
CC of the present sequence. The human kinases share structural
CC similarity with animal kinases, more particularly serine or
CC threonine protein kinases. Human kinase cDNA is useful for the
CC detection of mutant human kinase for the diagnosis of disease,
CC and also as a therapeutic. It is useful for screening drugs
CC effective in the treatment of symptomatic or phenotypic
CC manifestations perturbing the normal function of NHP in the
CC body. The NHP nucleotide sequences are useful for generation of
CC antibodies, as reagents in diagnostic assays, for the
CC identification of other cellular gene products related to human
CC kinases, and as reagents in assays for screening compounds that
CC are useful for treating mental, biological or medical disorders.
CC NHP oligonucleotides are used as probes. The labelled NHP probes
CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.
XX
SQ Sequence 1819 BP; 539 A; 368 C; 406 G; 505 T; 1 other;

Query Match 54.9%; Score 572; DB 22; Length 1819;
Best Local Similarity 100.0%; Pred. No. 2.9e-276;
Matches 572; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 454 attccaggagatgcctacacccgattatgtagctacgagatggtaccgagctcctgaaactt 513
|||||
Db 897 attccaggagatgcctacacccgattatgtagctacgagatggtaccgagctcctgaaactt 956
|||||
QY 514 cttgtgggagatactcagtagtggttcttcagtcgatataatgggctatgggtgtgttttt 573
|||||
Db 957 cttgtgggagatactcagtagtggttcttcagtcgatataatgggctatgggtgtgttttt 1016
|||||
QY 574 gcagagctcctgacagccagccactgtggcctggaaatcagatgtggaccaactttat 633
|||||
Db 1017 gcagagctcctgacagccagccactgtggcctggaaatcagatgtggaccaactttat 1076
|||||
QY 634 ctgataatcagaacactagtagagacggggttttcgccaatgttgaccaggctggtctcgaa 693
|||||
Db 1077 ctgataatcagaacactagtagagacggggttttcgccaatgttgaccaggctggtctcgaa 1136
|||||
QY 694 ctcttgacgtcaagtgatcacacctgcogtagcctctcaaatgtcgtggaatlacaggaaaa 753
|||||
Db 1137 ctcttgacgtcaagtgatcacacctgcogtagcctctcaaatgtcgtggaatlacaggaaaa 1196
|||||
QY 754 ttaatcccaagacatcaatcaatctttaaaaagtaacgggtttttccatggcatcagtata 813
|||||
Db 1197 ttaatcccaagacatcaatcaatctttaaaaagtaacgggtttttccatggcatcagtata 1256
|||||
QY 814 cctgagccagaagacatggaaactcttgaggaaaaagtcttcagatgttcacotgtggct 873
|||||
Db 1257 cctgagccagaagacatggaaactcttgaggaaaaagtcttcagatgttcacotgtggct 1316
|||||
QY 874 ctgaacttcocatgaaggggtgtctgaagatgaatccagatgacagattaacctgttcccaa 933
|||||
Db 1317 ctgaacttcocatgaaggggtgtctgaagatgaatccagatgacagattaacctgttcccaa 1376
|||||
QY 934 ctccctggagagctcctacttttgattcttttcaagaggcccaaatataaagaaaaagcacgt 993
|||||
Db 1377 ctccctggagagctcctacttttgattcttttcaagaggcccaaatataaagaaaaagcacgt 1436
|||||
QY 994 aatgaaggagaagaacagagacgccaacaggt 1025
|||||
Db 1437 aatgaaggagaagaacagagacgccaacaggt 1468
|||||

RESULT 6
AAD03812
ID AAD03812 standard; cDNA; 561 BP.
XX
AC AAD03812;
XX
DT 19-JUN-2001 (first entry)
XX
DE Human kinase cDNA #1.
XX
KW Human; kinase; gene therapy; bioreactor; mental disorder;
KW biological disorder; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..561
FT /*tag= a
FT /product= "Human kinase #1"
FT /note= "The coding region does not include stop codon"
FT /partial
XX
PN WO200123579-A1.
XX
PD 05-APR-2001.
XX
PF 27-SEP-2000; 2000WO-US26621.
XX
PR 28-SEP-1999; 99US-0156511.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX

PI Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
XX
DR WPI; 2001-266166/27.
DR P-PSDB; AAE00490.
XX
PT New isolated human kinase polynucleotide useful for generating
PT antibodies, as reagents in diagnostic assays and for screening for
PT compounds useful for treating mental, biological or medical diseases
XX
PS Disclosure; Page 27; 38pp; English.
XX
CC The present sequence is a cDNA encoding novel human protein
CC (NHP) known as human kinase. The human kinases share structural
CC similarity with animal kinases, more particularly serine or
CC threonine protein kinases. Human kinase cDNA is useful for the
CC detection of mutant human kinase for the diagnosis of disease,
CC and also as a therapeutic. It is useful for screening drugs
CC effective in the treatment of symptomatic or phenotypic
CC manifestations perturbing the normal function of NHP in the
CC body. The NHP nucleotide sequences are useful for generation of
CC antibodies, as reagents in diagnostic assays, for the
CC identification of other cellular gene products related to human
CC kinases, and as reagents in assays for screening compounds that
CC are useful for treating mental, biological or medical disorders.
CC NHP oligonucleotides are used as probes. The labelled NHP probes
CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.
XX
SQ Sequence 561 BP; 189 A; 87 C; 112 G; 173 T; 0 other;

Query Match 43.6%; Score 454; DB 22; Length 561;
Best Local Similarity 100.0%; Pred. No. 3.le-217;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atggaaaaagtatgaaaaattagctaagactggagaagggtcttatgggtgtattcaaa 60
|||||
Db 1 atggaaaaagtatgaaaaattagctaagactggagaagggtcttatgggtgtattcaaa 60
|||||
QY 61 tgcagaaaaacaaacctctggacaaagttagctgtttaaaaaatttgggaacttgaagat 120
|||||
Db 61 tgcagaaaaacaaacctctggacaaagttagctgtttaaaaaatttgggaacttgaagat 120
|||||
QY 121 gatcctgttgttaagaaaaatagcactaaagagaaaatcgtatgttgaagcaattaaaaacat 180
|||||
Db 121 gatcctgttgttaagaaaaatagcactaaagagaaaatcgtatgttgaagcaattaaaaacat 180
|||||
QY 181 ccaaatcttgtgaacctcatcgagggtgttcaggagaaaaaggaaaaatgcattagtttt 240
|||||
Db 181 ccaaatcttgtgaacctcatcgagggtgttcaggagaaaaaggaaaaatgcattagtttt 240
|||||
QY 241 gaatactgtgatcatcacactttttaaagtgcgtggaaagaaaccccaatggagttgctgat 300
|||||
Db 241 gaatactgtgatcatcacactttttaaagtgcgtggaaagaaaccccaatggagttgctgat 300
|||||
QY 301 ggagtgatcaaaaagcgtattatggcaaacacitcgaagcttcttaatttctgtcatacat 360
|||||
Db 301 ggagtgatcaaaaagcgtattatggcaaacacitcgaagcttcttaatttctgtcatacat 360
|||||
QY 361 aactgtattcacagagatataaaaacctgaaaaatattcttaactaagcaaggaataatc 420
|||||
Db 361 aactgtattcacagagagatataaaaacctgaaaaatattcttaactaagcaaggaataatc 420
|||||
QY 421 aagatttggacttcgggtttgcacaaattctga 454
|||||
Db 421 aagatttggacttcgggtttgcacaaattctga 454
|||||

PS Example 1; Figure 1; 433pp; English.

XX AAS06701-AAS06757 encode for novel human protein kinases #1-57. The

CC novel protein kinases have been identified as members of the tyrosine

CC or serine/threonine kinase (PTK and STK) families. The polynucleotides

CC encoding protein kinases and the polypeptides may be used in the

CC prevention, diagnosis and treatment of diseases associated with

CC inappropriate kinase expression. For example, they may be used to treat

CC cancers (especially cancers of haematopoietic origin), cardiovascular

CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),

CC immune related diseases (e.g. rheumatoid arthritis), neurological

CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.

CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious

CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).

CC Additionally, polynucleotides encoding protein kinases may be

CC used for gene therapy and as DNA probes in diagnostic assays.

CC The protein kinase polypeptides may be used as antigens in the production

CC of antibodies against the protein kinases and in assays to identify

CC modulators of protein kinase expression and activity.

XX Sequence 1083 BP; 366 A; 199 C; 225 G; 293 T; 0 other;

SQ

Query Match 43.6%; Score 454; DB 22; Length 1083;

Best Local Similarity 100.0%; Pred. No. 3.2e-217;

Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atggaaaagtatgaaaaattagctaaagactggagaagggtcttattggttgatttcaaa 60

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

1 atggaaaagtatgaaaaattagctaaagactggagaagggtcttattggttgatttcaaa 60

QY 61 tgcagaaacaaaacctctggacaagttagctgtgttataaaaaatttggaaatcgaagt 120

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

61 tgcagaaacaaaacctctggacaagttagctgtgttataaaaaatttggaaatcgaagt 120

QY 121 gatcctgtttaaagaaaatagcactaagagaatactatgttgaagcaattataaacat 180

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

121 gatcctgtttaaagaaaatagcactaagagaatactatgttgaagcaattataaacat 180

QY 181 ccaaatcttgtgaacctcatcgaggtgttcaggagaaaaaggaaaaatgcatttagtttt 240

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

181 ccaaatcttgtgaacctcatcgaggtgttcaggagaaaaaggaaaaatgcatttagtttt 240

QY 241 gaatactgtgatcatcacacttttaaatgagctggaaagaaaccccaaatggagttgctgat 300

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

241 gaatactgtgatcatcacacttttaaatgagctggaaagaaaccccaaatggagttgctgat 300

QY 301 ggagtgatcaaaagcgtattatggcaaacacttcaagctctttaatttctgtcatatacat 360

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

301 ggagtgatcaaaagcgtattatggcaaacacttcaagctctttaatttctgtcatatacat 360

QY 361 aactgtatttcacagagatataaaacctgaaaaatattctaactaagcaaggaataatc 420

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

361 aactgtatttcacagagatataaaacctgaaaaatattctaactaagcaaggaataatc 420

QY 421 aagatttgtgacttcgggtttgcacaaaattctga 454

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

421 aagatttgtgacttcgggtttgcacaaaattctga 454

RESULT 9

AAF44669

ID AAF44669 standard; cDNA; 911 BP.

XX

AC AAF44669;

XX

DT 27-MAR-2001 (first entry)

XX

DE Novel protein kinase cDNA, SEQ ID NO: 49.

XX

KW Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;

KW immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic;

KW dermatological; antidiabetic; antiinfertility; gene therapy; vaccine;

KW immune disorder; cardiovascular disease; neurodegenerative disease;

KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;

KW inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.

XX Homo sapiens.

OS

XX WO200073469-A2.

PN

XX 07-DEC-2000.

PD

XX 26-MAY-2000; 2000WO-US14842.

PF

XX 28-MAY-1999; 99US-0136503.

PR

XX (SUGE-) SUGEN INC.

PA

XX Plowman GD, Martinez R, Whyte D, Sudersanam S;

PI

XX WPI; 2001-032161/04.

DR

DR P-PSDB; AAB65642.

XX

PT Nucleic acids encoding kinase polypeptides, useful for diagnosing and

PT treating immune-related diseases and disorders, cardiovascular disease,

PT neurodegenerative diseases and/or cancers -

XX Example 4; Fig 2; 310pp; English.

PS

XX The present sequence encodes a novel protein kinase. The nucleic acids

CC and the protein kinases they encode may be used in the treatment and

CC diagnosis of diseases associated with inappropriate kinase expression

CC such as immune-related diseases and disorders, cardiovascular disease,

CC neurodegenerative diseases and/or cancers. The nucleic acids and

CC complementary sequences may also be used as DNA probes in diagnostic

CC assays. The kinase polypeptides may be used as antigens in the production

CC of antibodies of kinase expression and activity. Anti-kinase antibodies

CC and kinase antagonists may also be used to down regulate kinase

CC expression and activity. Diseases related to kinase expression and

CC activity include rheumatoid arthritis, atherosclerosis, autoimmune

CC disorders, complications of organ transplantation, myocardial infarction,

CC immune disorders, cardiomyopathies, strokes, renal failure,

CC oxidative-stress related disorders, chronic inflammatory bowel disease,

CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,

CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and

CC reproductive disorders.

XX

SQ Sequence 911 BP; 293 A; 182 C; 178 G; 258 T; 0 other;

Query Match 35.4%; Score 368; DB 22; Length 911;

Best Local Similarity 100.0%; Pred. No. 3.4e-174;

Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 285 aaatggagttgctgatggagtgatcaaaagcgtattatggcaaacacttcaagctcttaa 344

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

1 aaatggagttgctgatggagtgatcaaaagcgtattatggcaaacacttcaagctcttaa 60

QY 345 ttctgtcatatatacataaactgtatttcacagagatataaaacctgaaaaatattctaatac 404

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

61 ttctgtcatatatacataaactgtatttcacagagatataaaacctgaaaaatattctaatac 120

QY 405 taagcaaggaataaatcaagatttgcacttcgggtttgcacaaaattctgattccaggaga 464

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

121 taagcaaggaataaatcaagatttgcacttcgggtttgcacaaaattctgattccaggaga 180

QY 465 tgcctacaccgattatgtagctacgagatggtaccgagctcctgaactcttctgtgggaga 524

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

181 tgcctacaccgattatgtagctacgagatggtaccgagctcctgaactcttctgtgggaga 240

QY 525 tactcagtgatggttcttcagtcgatatatggcgctattggtgtgtttttgcagagctcct 584

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

241 tactcagtgatggttcttcagtcgatatatggcgctattggtgtgtttttgcagagctcct 300

QY 585 gacaggccagccactgtggcctggaaaaatcagatgtggaccaactttatctgataatcag 644

Db 301 gacagggcagccactgtggcciggaataatcagatgtggaccaacttatctgataatcag 360
QY 645 aacactag 652
Db 361 aacactag 368

RESULT 10
AAF97873/c
ID AAF97873 standard; DNA; 12729 BP.
XX AC AAF97873;
XX DT 31-MAY-2001 (first entry)
XX DE Human neuroblastoma cell line NB-1 lp36 nucleotide sequence SEQ ID NO:87.
XX DE
XX KW Human; chromosome 1; lp36; neuroblastoma cell line; NB-1; anticancer;
KW tumour suppressor; human lp36 homozygosity deletion domain; tumour;
KW diagnosis; ds.
XX OS Homo sapiens.
XX XX
XX PN WO200116311-A1.
XX PD 08-MAR-2001.
XX DX
XX PF 31-AUG-2000; 2000WO-JP05930.
XX XX
XX PR 31-AUG-1999; 99JP-0245962.
XX PR 09-MAY-2000; 2000JP-0136266.
XX XX
XX PA (HISM) HISAMITSU PHARM CO LTD.
XX PA (CHIB-) CHIBA PREFECTURE.
XX PI Nakagawara A;
XX XX
XX DR WPI; 2001-226686/23.
XX XX
XX PT Human lp36 homozygosity deletion domain from the 36-position of first
PT chromosome short arm in human neuroblastoma cell lines, applicable e.g.
PT in gene diagnosis of tumors as well as in developing anti-cancer drugs
PT -
XX XX
XX PS Example 8; Page 213-218; 226pp; Japanese.
XX XX
XX CC The present invention describes a homozygosity deletion domain
CC co-existing in the 36-position of the first chromosome short arm (lp36)
CC in human neuroblastoma. Also described are base sequences from the lp36
CC position of human neuroblastoma cell lines (NB-1 and MASS-NB-SCH-1),
CC which are tumour suppressor genes in human neuroblastoma. The genes are
CC tumour markers and reagents in studying mechanism of tumour body
CC formation, and gene diagnosis of tumours as well as in developing
CC anti-cancer drugs. AAF97787 to AAF97829 represent PCR primers used in
CC the exemplification of the present invention, and AAF97830 to AAF97874
CC represent sequences given in the exemplification of the present
XX invention.
XX SQ Sequence 12729 BP; 4419 A; 2482 C; 2733 G; 3095 T; 0 other;

Query Match 3.8%; Score 40; DB 22; Length 12729;
Best Local Similarity 100.0%; Pred. No. 4.7e-10;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacggggttcgccatgttgaccaggctgtct 689
Db 1800 TAGTAGAGACGGGTTTCGCCATGTTGACCAGGCTGGTCT 1761

RESULT 11

AAC29051
ID AAC29051 standard; cDNA; 105 BP.
XX AC AAC29051;
XX DT 06-OCT-2000 (first entry)
XX DE Human secreted protein 5' EST, SEQ ID NO: 33126.
XX DE
XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX OS Homo sapiens.
XX XX
XX PN EP1033401-A2.
XX PD 06-SEP-2000.
XX XX
XX PF 21-FEB-2000; 2000EP-0200610.
XX XX
XX PR 26-FEB-1999; 99US-0122487.
XX XX
XX PA (GEST) GENSET.
XX XX
XX PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX DX
XX DR WPI; 2000-500381/45.
XX XX
XX PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX PS Claim 1; SEQ ID 33126; 71pp + CD-ROM; English.
XX XX
XX CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
XX SQ Sequence 105 BP; 17 A; 27 C; 29 G; 32 T; 0 other;

Query Match 3.2%; Score 33; DB 21; Length 105;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 ccatgttgaccaggctggtctcgaactcttgac 701
Db 11 ccatgttgaccaggctggtctcgaactcttgac 43

RESULT 12
AAQ61041
ID AAQ61041 standard; DNA; 411 BP.
XX AC AAQ61041;
XX DT 16-MAR-1994 (first entry)
XX DE Human brain Expressed Sequence Tag EST01677.
XX DE
XX KW Gene transcription product; genetic markers; tagging; in vivo;
KW transcription; mapping; locations; chromosomes; chromosomal; ss.

XX Homo sapiens.
OS WO9316178-A.
PN
XX
XX
PD 19-AUG-1993.
XX
XX 12-FEB-1993; 93WO-US01294.
PF
XX 12-FEB-1992; 92US-0837195.
PR
XX (USSH) US DEPT HEALTH & HUMAN SERVICE.
PA
XX Adams MD, Moreno RF, Venter CJ;
PI
XX WPI; 1993-272882/34.
DR
XX Enriched oligonucleotides and corresp. sequences - used as
PT markers for human genes transcribed in-vivo, facilitate tagging
PT of most human genes
PT
XX
XX Example 4; Page 430; 500pp; English.
PS
XX The Expressed Sequence Tag was isolated from a human brain cDNA
CC library as part of a large set of ESTs which can be used as markers
CC for human genes transcribed in vivo. They can be used to facilitate
CC tagging of most human genes, for mapping locations of expressed genes
CC on chromosomes, for individual or forensic identification, for mapping
CC locations of disease-associated genes, for identification of tissue
CC type, and for prepn. of antisense sequences, probes and constructs.
CC EST01677 has a "poor" coding probability as evaluated using the
CC coding-region prediction program CRM. See also AAQ59041-Q61440.
XX
XX Sequence 411 BP; 99 A; 114 C; 93 G; 105 T; 0 other;
SQ

Query Match 3.2%; Score 33; DB 14; Length 411;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 ccattgtgaccaggctggtctcgaaactcttgac 701
Db 63 ccattgtgaccaggctggtctcgaaactcttgac 95
|||||

RESULT 13
AAF26586/c
ID AAF26586 standard; DNA; 2286 BP.
XX
XX AAF26586;
AC
XX 27-MAR-2001 (first entry)
DT
XX DNA encoding human secreted protein #40.
DE
XX Secreted protein; gene therapy; vaccine; cancer; leukemia;
KW autoimmune disease; allergy; inflammation; graft rejection;
KW hyperproliferation; cardiovascular; infection; ss.
XX
OS Homo sapiens.
XX WO200076531-A1.
PN
XX 21-DEC-2000.
PD
XX
XX 01-JUN-2000; 2000WO-US15137.
PF
XX 11-JUN-1999; 99US-0138625.
PR
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Rosen CA, Ruben SM, Komatsoulis GA;
PI
XX

DR WPI; 2001-071148/08.
XX
XX Nucleic acids encoding 47 human secreted polypeptides, useful for
PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
PT disease and diabetic retinopathy -
XX
XX Claim 1; Page 467-498; 525pp; English.
PS
XX The present invention relates to 26 secreted human proteins. The
CC proteins may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate polypeptide expression.
CC For example, they may be used in gene therapy or in vaccines.
CC Typical of diseases which are potentially treatable are cancers
CC (including leukemia), autoimmune diseases, allergies, inflammation,
CC graft rejection, hyperproliferation, cardiovascular diseases
CC (particularly critical limb ischemia and coronary disease) and any
CC involving abnormal angiogenesis, neurodegeneration and/or
CC infectious diseases.
XX
XX Sequence 2286 BP; 692 A; 422 C; 479 G; 692 T; 1 other;
SQ

Query Match 3.2%; Score 33; DB 22; Length 2286;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 664 ttctgccatgttgaccaggctggtctcgaaactc 696
Db 2111 TTTCGCCATCTTGACCAGGCTGGTCTCGAACTC 2079
|||||

RESULT 14
AAI60415/c
ID AAI60415 standard; cDNA; 5600 BP.
XX
XX AAI60415;
AC
XX 22-OCT-2001 (first entry)
DT
XX Human polynucleotide SEQ ID NO 4404.
DE
XX
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX WO200153312-A1.
PN
XX 26-JUL-2001.
PD
XX 26-DEC-2000; 2000WO-US34263.
PF
XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
DR P-PSDB; AAM41259.
XX

XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
XX Claim 1; SEQ ID NO 4404; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 5600 BP; 1708 A; 1046 C; 1174 G; 1672 T; 0 other;

Query Match 3.2%; Score 33; DB 22; Length 5600;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 664 ttctgccatgttgaccaggctggtctcgaaactc 696
|||||
Db 4303 TTTCGCCCATGTTGACCAGGCTGGTCTCGAACTC 4271
|||||

RESULT 15
AAI58629/c
ID AAI58629 standard; cDNA; 6027 BP.
XX
XX AAI58629;
XX
XX
DT 22-OCT-2001 (first entry)
XX
XX
DE Human polynucleotide SEQ ID NO 832.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
XX WO200153312-A1.
PN
XX
PD 26-JUL-2001.
XX
XX 26-DEC-2000; 2000WO-US34263.
PF
XX
XX 21-JAN-2000; 2000US-0488725.
PR
PR 25-APR-2000; 2000US-0552317.
PR
PR 09-JUL-2000; 2000US-0598042.
PR
PR 19-JUL-2000; 2000US-0620312.
PR
PR 03-AUG-2000; 2000US-0653450.
PR
PR 14-SEP-2000; 2000US-0662191.
PR
PR 19-OCT-2000; 2000US-0693036.
PR
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR P-PSDB; AAM39473.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
XX Claim 1; SEQ ID NO 832; 10078pp; English.
PS
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 6027 BP; 1786 A; 1180 C; 1301 G; 1760 T; 0 other;

Query Match 3.2%; Score 33; DB 22; Length 6027;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 664 ttctgccatgttgaccaggctggtctcgaaactc 696
|||||
Db 4643 TTTCGCCCATGTTGACCAGGCTGGTCTCGAACTC 4611
|||||

RESULT 16
AAI64117/c
ID AAI64117 standard; cDNA; 32191 BP.
XX
XX AAI64117;
XX
XX 24-OCT-2001 (first entry)
DT
XX
XX
DE Human bladder related polynucleotide, SEQ ID NO: 150.
XX
KW Human; bladder antigen; cytostatic; immunosuppressive; nootropic;
KW neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;
KW antiinflammatory; antiulcer; vulnerary; anticonvulsant; antibacterial;
KW antifungal; antiparasitic; cardiant; gene therapy; cancer;
KW immune disorder; cardiovascular disorder; wound healing; infection;
KW neurological disease; ss.
XX
OS Homo sapiens.
XX
XX WO200159064-A2.
PN
XX
PD 16-AUG-2001.
XX
XX
PF 17-JAN-2001; 2001WO-US01342.
XX
PR 31-JAN-2000; 2000US-0179065.
PR
PR 04-FEB-2000; 2000US-0180628.
PR
PR 24-FEB-2000; 2000US-0184664.
PR
PR 02-MAR-2000; 2000US-0186350.
PR
PR 16-MAR-2000; 2000US-0189874.
PR
PR 17-MAR-2000; 2000US-0190076.
PR
PR 18-APR-2000; 2000US-0198123.
PR
PR 19-MAY-2000; 2000US-0205515.
PR
PR 07-JUN-2000; 2000US-0209467.

us-09-671-050-9.oli20.rng

Fri May 3 11:02:10 2002

PR	28-JUN-2000;	2000US-0214886.	PR	20-OCT-2000;	2000US-0241786.
PR	30-JUN-2000;	2000US-0215135.	PR	20-OCT-2000;	2000US-0241787.
PR	07-JUL-2000;	2000US-0216647.	PR	20-OCT-2000;	2000US-0241808.
PR	07-JUL-2000;	2000US-0216880.	PR	20-OCT-2000;	2000US-0241809.
PR	11-JUL-2000;	2000US-0217487.	PR	20-OCT-2000;	2000US-0241826.
PR	11-JUL-2000;	2000US-0217496.	PR	01-NOV-2000;	2000US-0244617.
PR	14-JUL-2000;	2000US-0218290.	PR	08-NOV-2000;	2000US-0246474.
PR	26-JUL-2000;	2000US-0220963.	PR	08-NOV-2000;	2000US-0246475.
PR	26-JUL-2000;	2000US-0220964.	PR	08-NOV-2000;	2000US-0246476.
PR	14-AUG-2000;	2000US-0224518.	PR	08-NOV-2000;	2000US-0246477.
PR	14-AUG-2000;	2000US-0224519.	PR	08-NOV-2000;	2000US-0246478.
PR	14-AUG-2000;	2000US-0225213.	PR	08-NOV-2000;	2000US-0246523.
PR	14-AUG-2000;	2000US-0225214.	PR	08-NOV-2000;	2000US-0246524.
PR	14-AUG-2000;	2000US-0225266.	PR	08-NOV-2000;	2000US-0246525.
PR	14-AUG-2000;	2000US-0225267.	PR	08-NOV-2000;	2000US-0246526.
PR	14-AUG-2000;	2000US-0225268.	PR	08-NOV-2000;	2000US-0246527.
PR	14-AUG-2000;	2000US-0225270.	PR	08-NOV-2000;	2000US-0246528.
PR	14-AUG-2000;	2000US-0225447.	PR	08-NOV-2000;	2000US-0246532.
PR	14-AUG-2000;	2000US-0225757.	PR	08-NOV-2000;	2000US-0246609.
PR	14-AUG-2000;	2000US-0225758.	PR	08-NOV-2000;	2000US-0246610.
PR	14-AUG-2000;	2000US-0226279.	PR	08-NOV-2000;	2000US-0246611.
PR	18-AUG-2000;	2000US-0226681.	PR	08-NOV-2000;	2000US-0246613.
PR	22-AUG-2000;	2000US-0226868.	PR	08-NOV-2000;	2000US-0249207.
PR	22-AUG-2000;	2000US-0227182.	PR	17-NOV-2000;	2000US-0249208.
PR	23-AUG-2000;	2000US-0227009.	PR	17-NOV-2000;	2000US-0249209.
PR	30-AUG-2000;	2000US-0228924.	PR	17-NOV-2000;	2000US-0249210.
PR	01-SEP-2000;	2000US-0229287.	PR	17-NOV-2000;	2000US-0249211.
PR	01-SEP-2000;	2000US-0229343.	PR	17-NOV-2000;	2000US-0249212.
PR	01-SEP-2000;	2000US-0229344.	PR	17-NOV-2000;	2000US-0249213.
PR	01-SEP-2000;	2000US-0229345.	PR	17-NOV-2000;	2000US-0249214.
PR	05-SEP-2000;	2000US-0229509.	PR	17-NOV-2000;	2000US-0249215.
PR	05-SEP-2000;	2000US-0229513.	PR	17-NOV-2000;	2000US-0249216.
PR	06-SEP-2000;	2000US-0230437.	PR	17-NOV-2000;	2000US-0249217.
PR	06-SEP-2000;	2000US-0230438.	PR	17-NOV-2000;	2000US-0249218.
PR	08-SEP-2000;	2000US-0231242.	PR	17-NOV-2000;	2000US-0249244.
PR	08-SEP-2000;	2000US-0231243.	PR	17-NOV-2000;	2000US-0249264.
PR	08-SEP-2000;	2000US-0231244.	PR	17-NOV-2000;	2000US-0249265.
PR	08-SEP-2000;	2000US-0231413.	PR	17-NOV-2000;	2000US-0249297.
PR	08-SEP-2000;	2000US-0231414.	PR	17-NOV-2000;	2000US-0249299.
PR	08-SEP-2000;	2000US-0232080.	PR	17-NOV-2000;	2000US-0249300.
PR	08-SEP-2000;	2000US-0232081.	PR	01-DEC-2000;	2000US-0250160.
PR	12-SEP-2000;	2000US-0231968.	PR	01-DEC-2000;	2000US-0250391.
PR	14-SEP-2000;	2000US-0232397.	PR	05-DEC-2000;	2000US-0251030.
PR	14-SEP-2000;	2000US-0232398.	PR	05-DEC-2000;	2000US-0251988.
PR	14-SEP-2000;	2000US-0232399.	PR	05-DEC-2000;	2000US-0256719.
PR	14-SEP-2000;	2000US-0232400.	PR	06-DEC-2000;	2000US-0251479.
PR	14-SEP-2000;	2000US-0232401.	PR	08-DEC-2000;	2000US-0251856.
PR	14-SEP-2000;	2000US-0233063.	PR	08-DEC-2000;	2000US-0251868.
PR	14-SEP-2000;	2000US-0233064.	PR	08-DEC-2000;	2000US-0251869.
PR	14-SEP-2000;	2000US-0233065.	PR	08-DEC-2000;	2000US-0251989.
PR	21-SEP-2000;	2000US-0234223.	PR	08-DEC-2000;	2000US-0251990.
PR	21-SEP-2000;	2000US-0234274.	PR	11-DEC-2000;	2000US-0254097.
PR	25-SEP-2000;	2000US-0234997.	PR	05-JAN-2001;	2001US-0259678.
PR	25-SEP-2000;	2000US-0234998.	XX		
PR	26-SEP-2000;	2000US-0235484.	PA		
PR	27-SEP-2000;	2000US-0235834.	PI		
PR	27-SEP-2000;	2000US-0235836.	XX		
PR	29-SEP-2000;	2000US-0236327.	XX		
PR	29-SEP-2000;	2000US-0236367.	DR		
PR	29-SEP-2000;	2000US-0236368.	XX		
PR	29-SEP-2000;	2000US-0236369.	PT		
PR	29-SEP-2000;	2000US-0236370.	PT		
PR	02-OCT-2000;	2000US-0236802.	PT		
PR	02-OCT-2000;	2000US-0237037.	XX		
PR	02-OCT-2000;	2000US-0237038.	PS		
PR	02-OCT-2000;	2000US-0237039.	XX		
PR	02-OCT-2000;	2000US-0237040.	CC		
PR	13-OCT-2000;	2000US-0239935.	CC		
PR	13-OCT-2000;	2000US-0239937.	CC		
PR	20-OCT-2000;	2000US-0240960.	CC		
PR	20-OCT-2000;	2000US-0241221.	CC		
PR	20-OCT-2000;	2000US-0241785.			

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-514652/56.

Forty five bladder related polynucleotides, useful in the prevention, treatment and diagnosis of cancer, immune disorders, cardiovascular disorders and neurological diseases -

Disclosure; SEQ ID NO 150; 482pp + sequence listing; English.

The invention relates to forty five novel bladder related polynucleotides. The polynucleotides and the polypeptides that they encode are useful in the diagnosis, treatment and prevention of: cancer, particularly breast and ovarian cancer, and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital system; immune disorders such as

CC Addison's disease, allergies, autoimmune haemolytic anaemia,
CC autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple
CC sclerosis, rheumatoid arthritis and ulcerative colitis; cardiovascular
CC disorders such as myocardial ischaemias; wound healing; neurological
CC diseases such as cerebral anoxia and epilepsy; and infectious diseases
CC such as viral, bacterial, fungal and parasitic infections. Numerous
CC examples of each type of disorder are given in the specification.
CC The polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities. The polynucleotides are
CC useful for chromosome identification. They are also useful as probes
CC for diagnosing or treating a disorder related to the female reproductive
CC system, particularly breast and/or ovary cancer. The present
CC sequence is a bladder related polynucleotide encompassed by the
CC invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 32191 BP; 10213 A; 6968 C; 6580 G; 8430 T; 0 other;

Query Match 3.0%; Score 31; DB 22; Length 32191;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 660 ggggtttcccatgttgaccaggctggtctc 690
|||||
Db 18731 GGGGTTTCCCATGTTGACCAGGCTGGTCTC 18701

RESULT 17
AAX98426/c
ID AAX98426 standard; cDNA; 300 BP.

XX AAX98426;

DT 24-SEP-1999 (first entry)

DE Human cancer cell derived cDNA #152.

XX Cancer; human; colon; breast; lung; transmembrane receptor; ATPase;
KW integral membrane protein; aspartyl protease; GATA family; wnt family;
KW transcription factor; G-protein alpha subunit; protein phosphatase;
KW phorbol ester binding protein; diacylglycerol binding protein; trypsin;
KW protein kinase; tyrosine phosphatase; developmental signalling protein;
KW WW/rsp5/WWP domain; therapy; forensic; genetic mapping; diagnostic;
KW detection; treatment; cervical; melanoma; colorectal adenocarcinoma;
KW Wilm's tumour; retinoblastoma; sarcoma; myosarcoma; lung carcinoma;
KW leukemia; lymphoma; dysplasia; hyperplasia; endometrium; adrenal;
KW prostate; ss.

XX Homo sapiens.

XX WO9933982-A2.

XX 08-JUL-1999.

PF 22-DEC-1998; 98WO-US27610.

XX 21-DEC-1998; 98US-0217471.

PR 23-DEC-1997; 97US-0068755.

PR 03-APR-1998; 98US-0080664.

PR 21-OCT-1998; 98US-0105234.

PR 27-OCT-1998; 98US-0105877.

XX (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.

XX Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;

PI Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;

PI Jones LW, Kassam A, Kennedy GC, Kita D, Labat I;

PI Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;

PI Stache-Crain B, Sudduth-Klinger J, Williams LT;

XX
DR

WPI; 1999-430243/36.

XX
PT New isolated human polynucleotides

XX
PS Claim 1; Page 354; 591pp; English.

XX This invention describes novel isolated human polynucleotides obtained
CC by screening for differential expression in colon cancer, breast cancer
CC and lung cancer cell lines. The polynucleotides of the invention are
CC represented in AAX98275-X99118 and encode polypeptides of protein
CC families selected from 4 transmembrane segments integral membrane
CC proteins, 7 transmembrane receptors, ATPases associated with various
CC cellular activities (AAA), eukaryotic aspartyl proteases, GATA family of
CC transcription factors, G-protein alpha subunit, phorbol esters or
CC diacylglycerol binding proteins, protein kinase, protein phosphatase 2C,
CC protein tyrosine phosphatase, trypsin, wnt family of developmental
CC signalling proteins and WW/rsp5/WWP domain containing proteins. The
CC encoded polypeptides also have a functional domain selected from Ank
CC repeat, basic region plus leucine zipper transcription factors,
CC bromodomain, EF-hand, SH3 domain, WD domain/G-beta repeats, zinc finger
CC (C2H2 type), zinc finger (CCHC class), and zinc-binding metalloprotease
CC domain. The polynucleotides encode polypeptides with similarity to known
CC protein families and are predicted to have similar properties. The novel
CC polynucleotides can be used to develop products for use as therapeutic
CC agents and in forensics, genetic analysis, mapping and diagnostic
CC applications. In particular, the product can be used for the detection
CC and management of cancers. They can be used for treating e.g. cervical
CC cancers, melanomas, colorectal adenocarcinomas, Wilm's tumour, sarcomas,
CC retinoblastoma, myosarcomas, lung carcinomas, leukemias, such as chronic
CC myelogenous leukemia, promyelocytic leukemia, monocytic leukemia, and
CC myeloid leukemia, and lymphomas such as histiocytic lymphoma, anhydric
CC hereditary ectodermal dysplasia, congenital alveolar dysplasia,
CC epithelial dysplasia of the cervix, fibrous dysplasia of bone, and
CC mammary dysplasia, hyperplasias, e.g. endometrial, adrenal, breast,
CC prostate or thyroid hyperplasias or pseudoeptitheliomatous hyperplasia of
CC the skin.

XX
SQ Sequence 300 BP; 91 A; 58 C; 66 G; 85 T; 0 other;

Query Match 2.9%; Score 30; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 4.4e-05;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 660 ggggtttcccatgttgaccaggctggtctc 689
|||||

Db 249 GGGGTTTCCCATGTTGACCAGGCTGGTCT 220

RESULT 18
AAX98945/c

ID AAX98945 standard; cDNA; 752 BP.

XX AAX98945;

XX 24-SEP-1999 (first entry)

DT Human validated cancer cell derived cDNA #267.

DE Cancer; human; colon; breast; lung; transmembrane receptor; ATPase;
XX integral membrane protein; aspartyl protease; GATA family; wnt family;
KW transcription factor; G-protein alpha subunit; protein phosphatase;
KW phorbol ester binding protein; diacylglycerol binding protein; trypsin;
KW protein kinase; tyrosine phosphatase; developmental signalling protein;
KW WW/rsp5/WWP domain; therapy; forensic; genetic mapping; diagnostic;
KW detection; treatment; cervical; melanoma; colorectal adenocarcinoma;
KW Wilm's tumour; retinoblastoma; sarcoma; myosarcoma; lung carcinoma;
KW leukemia; lymphoma; dysplasia; hyperplasia; endometrium; adrenal;
KW prostate; ss.

XX Homo sapiens.

XX

AC AAH72939;
XX
DT 19-SEP-2001 (first entry)
XX
DE Human cervical cancer marker nucleic acid 4213.
XX
KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200142467-A2.
XX
PD 14-JUN-2001.
XX
XX 08-DEC-2000; 2000WO-US33312.
PF 08-DEC-1999; 99US-0169681.
XX 21-DEC-1999; 99US-0171350.
PR 14-MAR-2000; 2000US-0189315.
PR 12-MAY-2000; 2000US-0203791.
PR 09-JUN-2000; 2000US-0210600.
PR 21-JUL-2000; 2000US-0220114.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX
PI Schlegel R, Deeds J, Berger A, Zhao X;
XX
DR WPI; 2001-375006/39.
XX
PT New isolated nucleic acid for diagnosing and treating cervical cancer
PT and for assessing and detecting compounds for treating the cancer -
XX
XX
PS Claim 1; Page 922; 1051pp; English.
XX
CC The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
XX useful for gene therapy.
SQ Sequence 1718 BP; 516 A; 407 C; 303 G; 482 T; 10 other;

Query Match 2.9%; Score 30; DB 22; Length 1718;
Best Local Similarity 100.0%; Pred. No. 4.6e-05;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 661 ggggttcgccatgttgaccaggctggtctc 690
Db 19 ggggttcgccatgttgaccaggctggtctc 48

RESULT 21
AAS08146/c
ID AAS08146 standard; cDNA; 1762 BP.
XX
AC AAS08146;
XX
DT 23-OCT-2001 (first entry)
XX
DE Mammalian toxicological response marker #35.
XX
KW Mammalian toxicological response marker; antigen; antibody; agonist; ss.
XX
OS Homo sapiens.
XX
PN WO200136684-A2.
XX
PD 25-MAY-2001.
XX

PF 16-NOV-2000; 2000WO-US31743.
XX
PR 19-NOV-1999; 99US-0443184.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Cunningham MJ, Zweiger GB, Kaser MR, Panzer SR, Seilhamer JJ;
PI Yue H, Baughn MR, Azimzai Y, Lal P;
XX
DR WPI; 2001-355646/37.
XX
PT Novel mammalian nucleic acid molecules whose levels are up regulated or
PT down regulated following treatment with a toxic compound, useful for
PT detecting metabolic and toxicological responses and in monitoring drug
PT action -
XX
PS Claim 1; Page 48; 89pp; English.
XX
CC The sequence is a novel mammalian nucleic acid molecule whose levels are
CC up regulated or down regulated following treatment with a toxic compound.
CC Polynucleotide sequences complementary to the sequences of the invention
CC are useful for preventing a toxicological response by acting against one
CC or more up-regulated nucleic acid molecules. An agonist identified by the
CC above method is also useful for preventing a toxicological response by the
CC initiating transcription of a gene comprising a down regulated nucleic
CC acid molecule of the invention. Proteins encoded by the nucleic acids of
CC the invention are useful for producing antibodies. The nucleic acid
CC molecules are useful for producing an animal model system. The nucleic
CC acids immobilised on a substrate as hybridisable array element in a
CC microarray format may be used to characterise gene expression patterns
CC associated with novel compounds to elucidate any toxicological responses,
CC or to monitor the effects of treatments during clinical trials or therapy
CC where metabolic response to toxic compounds may be expected.
CC The nucleic acids are useful for various hybridisation technologies and
CC is useful for designing hybridisation probes. The nucleic acid molecule
CC or its fragment, or a protein encoded by the nucleic acid molecule may be
XX used to purify a ligand from a sample.
SQ Sequence 1762 BP; 497 A; 353 C; 443 G; 469 T; 0 other;

Query Match 2.9%; Score 30; DB 22; Length 1762;
Best Local Similarity 100.0%; Pred. No. 4.6e-05;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 660 ggggttcgccatgttgaccaggctggtctc 689
Db 1585 GGGGTTTCGCCATGTTGACCAGGCTGGTCT 1556

RESULT 22
AAH17495
ID AAH17495 standard; cDNA; 1804 BP.
XX
AC AAH17495;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:16962.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-0116126.
PF 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.

Fri May 3 11:02:10 2002

PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 8; SEQ ID 16962; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesising 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
XX Sequence 1804 BP; 440 A; 453 C; 470 G; 441 T; 0 other;
SQ

Query Match 2.9%; Score 30; DB 22; Length 1804;
Best Local Similarity 100.0%; Pred. No. 4.6e-05;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 672 tgttgaccaggctggtctcgaaactcttgac 701
|||||
Db 1258 tgttgaccaggctggtctcgaaactcttgac 1287
|||||

RESULT 23
AAH18014/c
ID AAH18014 standard; cDNA; 1916 BP.
XX
XX AAH18014;
AC
XX
XX 26-JUN-2001 (first entry)
DT
XX
XX Human cDNA sequence SEQ ID NO:17826.
DE
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
KW
XX Homo sapiens.
OS
XX EP1074617-A2.
PN
XX
XX 07-FEB-2001.
PD
XX
XX 28-JUL-2000; 2000EP-0116126.
PF
XX
XX 29-JUL-1999; 99JP-0248036.
PR

PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
PA
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI; 2001-318749/34.
DR
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 8; SEQ ID 17826; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesising 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
XX Sequence 1916 BP; 566 A; 352 C; 405 G; 593 T; 0 other;
SQ

Query Match 2.9%; Score 30; DB 22; Length 1916;
Best Local Similarity 100.0%; Pred. No. 4.6e-05;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 660 ggggtttcgccatgttgaccaggctggtct 689
|||||
Db 1314 GGGGTTTCGCCATGTTGACAGGCTGGTCT 1285
|||||

RESULT 24
AAI63723
ID AAI63723 standard; DNA; 9439 BP.
XX
XX AAI63723;
AC
XX
XX 22-OCT-2001 (first entry)
DT
XX
XX Human kidney related polynucleotide SEQ ID NO 1038.
DE
XX
XX Human; kidney antigen; immunosuppressive; antiarthritic; antirheumatic;
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
KW nootropic; neuroprotective; antibacterial; virucide; fungicide;
KW ophthalmological; antiallergic; hepatotropic; antidiabetic;
KW antiinflammatory; antiulcer; vulnery; anticonvulsant; antiparasitic;
KW gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; ds.
XX
XX Homo sapiens.
OS

PPI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-488784/53.
DR
XX New isolated nucleic acids and polypeptides, useful for diagnosing,
PT treating and/or preventing human diseases and disorders -
PT
XX Disclosure; SEQ ID NO 1038; 564pp + Sequence Listing; English.
PS
XX The invention relates to novel kidney related polynucleotides
CC (AAI62971-AAI63793) and the encoded polypeptides (AAM42417-AAM42691)
CC collectively known as kidney antigens and the use of such kidney antigens
CC for detecting disorders of the kidney, especially kidney cancer and
CC kidney cancer metastases. The polynucleotides and proteins are also
CC useful for preventing, treating or ameliorating medical conditions
CC e.g. by protein or gene therapy. The genes are isolated from a range
CC of human tissues disclosed in the specification. The nucleic acids,
CC proteins, antibodies and (ant)agonists are useful in the diagnosis,
CC treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer,
CC and other cancers of the adrenal gland, bone, bone marrow, breast,
CC gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders
CC e.g. Addison's disease, allergies, autoimmune haemolytic anaemia,
CC autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple
CC sclerosis, rheumatoid arthritis and ulcerative colitis;
CC (c) cardiovascular disorders such as myocardial ischaemias; (d) wound
CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
CC and (f) infectious diseases such as viral, bacterial, fungal and
CC parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WHO at ftp who.int/pub/published pct.sequences.

```

Query Match      2.8%; Score 29; DB 22; Length 9439;
Best Local Similarity 100.0%; pred. No. 0.00015;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 668 gccatgttgaccaggctggctcgaactc 696
|||||
pb 3769 gccatgttgaccaggctggctcgaactc 3797

RESULT 25
AAV57903/C
ID AAV57903 standard: DNA: 237326 BP.

AAV57903;	
21-DEC-1998	(first entry)
Hereditary haemochromatosis	subregion from an HH affected individual.
Bovine butyrophilin; BT;	human hereditary haemochromatosis; HFE;
diagnosis; iron metabolism; NPT3;	NPT4; RoRet; BTf1; BTf3;
BTf4; BTf5;	milk protein; lupus; Sjogren's syndrome; hypophosphatemia;
type 1 sodium transport gene;	ss.

PI Tsuchihashi Z, Wolff RK;
XX
XX WPI: 1998-240014/21.
XX Hereditary haemochromatosis gene products - used to develop products
PT for the diagnosis and treatment of hereditary disorders in iron
PT metabolism
XX
XX Claim 1; Fig 9; 209pp; English.
PS
XX The present invention describes hereditary haemochromatosis gene
CC products from the human haemochromatosis gene. The present sequence
CC represents a hereditary haemochromatosis subregion from an hereditary
CC haemochromatosis (HH) affected individual. Also described is a
CC method to determine the presence or absence of the common hereditary
CC haemochromatosis (HFE) gene mutation in an individual comprising:
CC (a) providing DNA or RNA from the individual; and (b) assessing the
CC DNA or RNA for the presence or absence of a haplotype or genotype where
CC the presence or absence of the haplotype genotype indicates the likely
CC presence of the HFE gene mutation in the genome of the individual. The
CC HFE gene sequences from the present invention can be used to develop
CC products for use in the diagnosis and treatment of HFE. The present
CC invention also describes BTf genes, which are homologues of the milk
CC protein butyrophilin (BTf), and can be used in the production of agonists
CC and antagonists of BTf function. Also described are: (1) a RoRet gene
CC which can be used to develop products for the study, diagnosis and
CC treatment of lupus and Sjogren's syndrome; and (2) NPT3 and NPT4 genes
CC which are homologues of a type 1 sodium transport gene, and can
CC similarly be used for hypophosphatemia.
XX sequence 237326 BP: 69596 A; 48904 C; 48217 G; 70609 T; 0 other;
XX

```

Query Match          2.8%; Score 29; DB 19; Length 237326;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 662 ggtttcgccatgttgaccaggctggtctc 690
| | | | | | | | | | | | | | |
pb 128378 ccttttgcacatgttgaccaggctggtctc 138301

RESULT 26
AAC05941
ID AAC05941 standard: CDNA: 261 BP.

AA	AAC05941;
AC	
XX	06-OCT-2000 (first entry)
DT	
XX	Human secreted protein 5' EST, SEQ ID NO: 10016.
DE	
XX	Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW	gene therapy; chromosome mapping; ss.
WU	

XX
DR
XX
PT
PT

WPI; 2000-500381/45.

New nucleic acid that is a 5' expressed sequence tag (5' EST) for obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for

PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX Claim 1; SEQ ID 10016; 71pp + CD-ROM; English.
PS
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX Sequence 261 BP; 62 A; 52 C; 55 G; 90 T; 2 other;
SQ

Query Match 2.7%; Score 28; DB 21; Length 261;
Best Local Similarity 100.0%; Pred. No. 0.00044;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 650 tagtagagacggggttcgccatgttga 677
Db 13 tagtagagacggggttcgccatgttga 40
|||||

RESULT 27
AAC05857
ID AAC05857 standard; cDNA; 323 BP.
XX
AC AAC05857;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 9932.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX WPI; 2000-500381/45.
XX
PS New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX Claim 1; SEQ ID 9932; 71pp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences

CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX Sequence 323 BP; 46 A; 86 C; 80 G; 110 T; 1 other;
SQ

Query Match 2.7%; Score 28; DB 21; Length 323;
Best Local Similarity 100.0%; Pred. No. 0.00045;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 669 ccattgtgaccaggctggtctcgaaactc 696
Db 248 ccattgtgaccaggctggtctcgaaactc 275
|||||

RESULT 28
AAH13269
ID AAH13269 standard; cDNA; 531 BP.
XX
AC AAH13269;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA clone (3'-primer) SEQ ID NO:10104.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
DR
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 3; SEQ ID 10104; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesising 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and

CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX SQ Sequence 531 BP; 112 A; 147 C; 121 G; 144 T; 7 other;

Query Match 2.7%; Score 28; DB 22; Length 531;
Best Local Similarity 100.0%; Pred. No. 0.00045;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 ccatgttgaccaggctggtctcgaactc 696
Db 175 ccatgttgaccaggctggtctcgaactc 202
|||||

RESULT 29
AAAL6349
ID AAAL6349 standard; DNA; 631 BP.
XX
AC AAAL6349;
XX
DT 14-JUN-2000 (first entry)
XX
DE Human colon cancer differentially expressed nucleotide sequence #354.
XX
KW Colon cancer; detect; differential expression; human; treatment;
KW detect mutation; non-invasive diagnostic method; ds.
XX
OS Homo sapiens.
XX WO200012702-A2.
PN
XX 09-MAR-2000.
PD
XX 30-AUG-1999; 99WO-US19424.
PF
XX 31-AUG-1998; 98US-0098639.
PR 27-JAN-1999; 99US-0117393.
XX
PA (FARB) BAYER CORP.
XX
PI Endege WO, Steinmann KE, Astle JH, Burgess CC, Carroll E;
PI Catino TJ, Dwivedi P, Ford DM, Lewis ME, Molino GA, Monahan JE;
PI Schlegel R;
XX
WPI; 2000-256641/22.
DR
XX Novel nucleic acids and proteins for identifying therapeutic agents -
PT useful for treating and diagnosing cancer, especially colon cancer -
XX
PS Claim 16; Page 267; 345pp; English.
XX
CC This sequence represents a human nucleotide sequence which is
CC differentially expressed in colon cancer cells compared to the expression
CC levels in normal cells. The nucleotide sequence can be used as a source
CC of primers and probes. The nucleotide sequence is useful for determining
CC the phenotype of a cell by detecting the differential expression of the
CC sequence relative to a normal cell. The probes derived from the sequence
CC can also be used to determine the phenotype of cells in a sample. Probes
CC and antibodies which hybridise to the nucleotide sequence can also be
CC used to determine the phenotype of a cell. The primers are useful for
CC detecting a mutation in a test nucleotide sequence and also for detecting
CC cancer, preferably colon cancer. Antibodies against the protein encoded
CC by the nucleotide sequence can also be used in a method to detect colon
CC cancer. The diagnostic method is non-invasive and accurate for diagnosing
CC colon cancer at an early stage.

XX SQ Sequence 631 BP; 124 A; 162 C; 154 G; 171 T; 20 other;

Query Match 2.7%; Score 28; DB 21; Length 631;
Best Local Similarity 100.0%; Pred. No. 0.00045;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 ccatgttgaccaggctggtctcgaactc 696
Db 214 ccatgttgaccaggctggtctcgaactc 241
|||||

RESULT 30
AAH04221
ID AAH04221 standard; cDNA; 805 BP.
XX
AC AAH04221;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA clone (5'-primer) SEQ ID NO:1056.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX Homo sapiens.
OS
PN EP1074617-A2.
XX
PD 07-FEB-2001.
PF
XX 28-JUL-2000; 2000EP-0116126.
PR
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 1; SEQ ID 1056; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesising 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632

CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 805 BP; 203 A; 200 C; 188 G; 207 T; 7 other;

Query Match 2.7%; Score 28; DB 22; Length 805;
Best Local Similarity 100.0%; Pred. No. 0.00045;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 663 gtttcgccatgttgaccaggctgtctc 690
|||||
Db 200 gtttcgccatgttgaccaggctgtctc 227

RESULT 31
AAH51279/c
ID AAH51279 standard; DNA; 1001 BP.
XX
AC AAH51279;
XX
DT 29-AUG-2001 (first entry)
XX
DE Human GSHS related DNA containing a biallelic polymorphism SEQ ID 170.
XX
KW Human; biallelic marker; single nucleotide polymorphism; SNP; MGSTII;
KW microsomal glutathione S-transferase II; malate decarboxylase enzyme;
KW DME1; ME1; cytochrome P450; glutathione reductase; GSHR; GSHS; GGT5;
KW flavin-containing monooxygenase; FMO; gamma-glutamyltransferase 5;
KW dipeptidase; DP; glucose 6-phosphate dehydrogenase; G6PDH; haplotype;
KW phosphogluconate dehydrogenase; PGDH; drug metabolism; phenotype;
KW uridine diphosphate glucoronosyl transferase; UGT2; asthma; hepatotoxicity;
KW zileuton; ds.
XX
OS Homo sapiens.
XX
PN WO200058508-A2.
XX
PD 05-OCT-2000.
XX
PF 24-MAR-2000; 2000WO-IB00403.
XX
PR 25-MAR-1999; 99US-0126269.
PR 30-APR-1999; 99US-0131961.
XX
PA (GEST) GENSET.
XX
XX Blumenfeld M, Bougueleret L, Chumakov I, Cohen-Akenine A;
PI
DR WPI; 2000-638353/61.
XX
PT Polynucleotides comprising sequences from malate decarboxylase
PT enzyme-related biallelic markers used for genotyping -
XX
PS Claim 13; Page 375-376; 673pp; English.
XX

Sequences AAH51110-AAH51593 represent human DNA fragments which contain
CC biallelic markers. The sequences are related to various human genes
CC including microsomal glutathione S-transferase II (MGSTII), malate
CC decarboxylase enzyme (DME1/ME1), cytochrome P450, glutathione
CC reductase/synthase (GSHR/GSHS), flavin-containing monooxygenases (FMO),
CC gamma-glutamyltransferase 5 (GGT5), dipeptidase (DP), glucose 6-phosphate
CC dehydrogenase (G6PDH), phosphogluconate dehydrogenase (PGDH), and uridine
CC diphosphate glucoronosyl transferases (UGT2). Each of these sequences
CC contains a biallelic marker/polymorphism, which is represented in the
CC sequence as a degenerate/undefined base. The genes to which the biallelic
CC marker containing sequences are related are involved in drug metabolism.
CC Sequences AAH51594 - AAH51598 represent the genomic sequence of the
CC MGSTII gene and four alternative MGSTII cDNA sequences. AAB62905-AAB62906
CC are MGSTII gene products. PCR primers AAH51599 and AAH51600 are used in
CC an example for the amplification of human genomic DNA fragments. The
CC invention includes a method of genotyping comprising determining the
CC identity of a nucleotide at a DME- or MGSTII-related biallelic marker in

CC a biological sample. The method is used to determine the frequency in
CC population of an allele of a DME- or MGST-II related biallelic marker and
CC to select an individual for inclusion in a clinical trial of a drug
CC treatment. The method is also used to detect association between allele
CC and phenotype, and to detect association between haplotype and phenotype.
CC The polynucleotides are used, in hybridization assays, sequencing assays
CC or allele specific amplification assays. The method can be used to
CC determine whether an individual suffers or is at risk of developing
CC asthma or is at risk of developing hepatotoxicity on treatment with
CC zileuton.
XX

SQ Sequence 1001 BP; 340 A; 200 C; 211 G; 249 T; 1 other;

Query Match 2.7%; Score 28; DB 21; Length 1001;
Best Local Similarity 100.0%; Pred. No. 0.00045;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 669 ccattgttgaccaggctgtctcgaactc 696
|||||
Db 745 CCATGTTGACCCAGGCTGGTCTCGAACTC 718

RESULT 32
AAH51559/c
ID AAH51559 standard; DNA; 1001 BP.
XX
AC AAH51559;
XX
DT 29-AUG-2001 (first entry)
XX
DE Human GLCL related DNA containing a biallelic polymorphism SEQ ID 450.
XX
KW Human; biallelic marker; single nucleotide polymorphism; SNP; MGSTII;
KW microsomal glutathione S-transferase II; malate decarboxylase enzyme;
KW DME1; ME1; cytochrome P450; glutathione reductase; GSHR; GSHS; GGT5;
KW flavin-containing monooxygenase; FMO; gamma-glutamyltransferase 5;
KW dipeptidase; DP; glucose 6-phosphate dehydrogenase; G6PDH; haplotype;
KW phosphogluconate dehydrogenase; PGDH; drug metabolism; phenotype;
KW uridine diphosphate glucoronosyl transferase; UGT2; asthma; hepatotoxicity;
KW zileuton; ds.
XX
OS Homo sapiens.
XX
PN WO200058508-A2.
XX
PD 05-OCT-2000.
XX
PF 24-MAR-2000; 2000WO-IB00403.
XX
PR 25-MAR-1999; 99US-0126269.
PR 30-APR-1999; 99US-0131961.
XX
PA (GEST) GENSET.
XX

Blumenfeld M, Bougueleret L, Chumakov I, Cohen-Akenine A;
PI
XX
DR WPI; 2000-638353/61.
XX
PT Polynucleotides comprising sequences from malate decarboxylase
PT enzyme-related biallelic markers used for genotyping -
XX
PS Claim 13; Page 624-625; 673pp; English.
XX

Sequences AAH51110-AAH51593 represent human DNA fragments which contain
CC biallelic markers. The sequences are related to various human genes
CC including microsomal glutathione S-transferase II (MGSTII), malate
CC decarboxylase enzyme (DME1/ME1), cytochrome P450, glutathione
CC reductase/synthase (GSHR/GSHS), flavin-containing monooxygenases (FMO),
CC gamma-glutamyltransferase 5 (GGT5), dipeptidase (DP), glucose 6-phosphate
CC dehydrogenase (G6PDH), phosphogluconate dehydrogenase (PGDH), and uridine
CC diphosphate glucoronosyl transferases (UGT2). Each of these sequences
CC contains a biallelic marker/polymorphism, which is represented in the

CC sequence as a degenerate/undefined base. The genes to which the biallelic
CC marker containing sequences are related are involved in drug metabolism.
CC Sequences AAH51594 - AAH51598 represent the genomic sequence of the
CC MGSTII gene and four alternative MGSTII cDNA sequences. AAB62905-AAB62906
CC are MGSTII gene products. PCR primers AAH51599 and AAH51600 are used in
CC an example for the amplification of human genomic DNA fragments. The
CC invention includes a method of genotyping comprising determining the
CC identity of a nucleotide at a DME- or MGSTII-related biallelic marker in
CC a biological sample. The method is used to determine the frequency in
CC population of an allele of a DME- or MGST-II related biallelic marker and
CC to select an individual for inclusion in a clinical trial of a drug
CC treatment. The method is also used to detect association between allele
CC and phenotype, and to detect association between haplotype and phenotype.
CC The polynucleotides are used, in hybridization assays, sequencing assays
CC or allele specific amplification assays. The method can be used to
CC determine whether an individual suffers or is at risk of developing
CC asthma or is at risk of developing hepatotoxicity on treatment with
CC zileuton.
XX
SQ Sequence 1001 BP; 329 A; 210 C; 214 G; 248 T; 0 other;

Query Match 2.7%; Score 28; DB 21; Length 1001;
Best Local Similarity 100.0%; Pred. No. 0.00045;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 ccatgttgaccaggctggtctcgaaactc 696
|||||
Db 788 CCATGTTGACCAAGCTGGTCTCGAACTC 761

RESULT 33
AAF85366
ID AAF85366 standard; cDNA; 1469 BP.
XX
AC AAF85366;
XX
DT 23-JUL-2001 (first entry)
XX
DE Nucleotide sequence of a human chemokine-alpha5.
XX
KW Bonzo; CXC chemokine receptor; inflammatory disease; cancer; infection;
KW SExCkine; spleen extracted chemokine; chemokine-alpha5; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 2..1306
FT /*tag= a
FT /product= "chemokine-alpha5"

XX WO200137872-A1.
PN
XX
PD 31-MAY-2001.
XX
PF 22-NOV-2000; 2000WO-US32206.
XX
PR 24-NOV-1999; 99US-0449437.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Briskin MJ, Murphy KE, Wilbanks AM, Wu L;
XX
DR WPI; 2001-343947/36.
DR P-PSDB; AAB68424.
XX
PT Identifying agents (especially antibodies) which bind to the CXC
PT chemokine receptor Bonzo, and which may be used to treat e.g. cancers
PT and inflammation -
XX
PS Disclosure; Fig 4A-C; 191pp; English.
XX
CC The present sequence encodes a human chemokine-alpha5. The specification

CC describes human Bonzo polypeptides. Bonzo is a CXC chemokine receptor.
CC The specification describes a method for identifying agents (especially
CC antibodies) which bind to Bonzo and inhibit the binding of a ligand
CC (especially SExCkine (spleen extracted chemokine)) and the agents
CC per se. The agents identified may be used for the treatment of a
CC disorder/disease related to aberrant Bonzo expression and activity,
CC such as inflammatory disease, cancers and/or infections (e.g. viral,
CC bacterial and fungal infections).
XX
SQ Sequence 1469 BP; 335 A; 427 C; 303 G; 404 T; 0 other;

Query Match 2.7%; Score 28; DB 22; Length 1469;
Best Local Similarity 100.0%; Pred. No. 0.00046;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 ccatgttgaccaggctggtctcgaaactc 696
|||||
Db 942 ccatgttgaccaggctggtctcgaaactc 969

RESULT 34
AAZ56728
ID AAZ56728 standard; cDNA; 1562 BP.
XX
AC AAZ56728;
XX
DT 23-MAR-2000 (first entry)
XX
DE Human transmembrane protein HTMPN-31 encoding cDNA.
XX
KW Human; transmembrane protein; HTMPN; diagnosis; immunospecific;
KW antiproliferative; neuroprotective; immune disorder;
KW reproductive disorder; smooth muscle disorder; neurological disorder;
KW gastrointestinal disorder; developmental disorder;
KW cell proliferative disorder; ss.
XX
OS Homo sapiens.
XX
PN WO9961471-A2.
XX
PD 02-DEC-1999.
XX
PF 28-MAY-1999; 99WO-US11904.
XX
PR 29-MAY-1998; 98US-0087260.
PR 02-JUL-1998; 98US-0091674.
PR 02-OCT-1998; 98US-0102954.
PR 24-NOV-1998; 98US-0109869.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Tang YT, Lal P, Hillman JL, Yue H, Guegler KJ, Corley NC;
PI Bandman O, Patterson C, Gorgone GA, Kaser MR, Baughn MR;
PI Au-Young J;
XX
DR WPI; 2000-072605/06.
DR P-PSDB; AAY57907.
XX
PT Proteins, polynucleotides, vectors, host cells and antibodies used to
PT diagnose, treat or prevent immune, reproductive, smooth muscle,
PT neurological, gastrointestinal, developmental and cell proliferative
PT disorders -
XX
PS Claim 9; Page 197; 229pp; English.
XX
CC AAZ56698 to AAZ56776 encode AAY57877 to AAY57955 which represent human
CC transmembrane proteins designated HTMPN-1 to HTMPN-79, respectively.
CC The transmembrane protein have immunospecific, antiproliferative and
CC neuroprotective activities. The human transmembrane proteins,
CC polynucleotides encoding them and other compositions and methods from
CC the present invention, can be used for the diagnosis, treatment or
CC prevention of immune, reproductive, smooth muscle, neurological,

CC gastrointestinal, developmental and cell proliferative disorders. The
CC HTPN's can be used to treat or prevent disorders associated with a
CC decreased expression or activity of HTPN.
XX
SQ Sequence 1562 BP; 324 A; 464 C; 404 G; 370 T; 0 other;

Query Match 2.7%; Score 28; DB 21; Length 1562;
Best Local Similarity 100.0%; Pred. No. 0.00046;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 ccatgttgaccaggctggtctcgaactc 696
|||||
Db 1459 ccatgttgaccaggctggtctcgaactc 1486
|||||

RESULT 35
AAH98793
ID AAH98793 standard; cDNA; 1624 BP.
XX
AC AAH98793;

XX 12-OCT-2001 (first entry)
XX Human EST-derived coding sequence SEQ ID NO: 650.

XX Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW diagnostics; forensic test; gene mapping; genetic disorder;
KW biodiversity; gene therapy; nutrition; ss.

XX Homo sapiens.
OS
XX WO200154477-A2.
PN
XX 02-AUG-2001.
PD
XX 25-JAN-2001; 2001WO-US02687.
PF
XX 25-JAN-2000; 2000US-0491404.
PR
17-JUL-2000; 2000US-0617746.
PR
03-AUG-2000; 2000US-0631451.
PR
15-SEP-2000; 2000US-0663870.

XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
PI Cao Y, Drmanac RA, Zhang J, Werhman T;
PI
XX WPI; 2001-476164/51.
DR
DR P-PSDB; AAM24134.

XX Isolated polypeptide for treatment of diseases, diagnostics, raising
PT antibodies and research use -
PT
XX Claim 1; Page 615; 1275pp; English.

XX The present invention provides the protein and coding sequences of novel
CC proteins from a variety of organisms, including human, dog, cat, horse,
CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
CC from the organism of interest. They can be used in diagnostics,
CC forensics, gene mapping, identification of mutations, to assess
CC biodiversity and for nutritional purposes. The present sequence is a cDNA
CC of the invention.

XX Sequence 1624 BP; 426 A; 347 C; 343 G; 507 T; 1 other;

Query Match 2.7%; Score 28; DB 22; Length 1624;
Best Local Similarity 100.0%; Pred. No. 0.00046;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 ccatgttgaccaggctggtctcgaactc 696
|||||
Db 312 ccatgttgaccaggctggtctcgaactc 339
|||||

RESULT 36
AAF90192
ID AAF90192 standard; cDNA; 1763 BP.
XX
AC AAF90192;

XX 06-AUG-2001 (first entry)

XX Nucleotide sequence of human spleen extracted chemokine (SExCkine).

XX G protein-coupled receptor ligand; spleen extracted chemokine; SExCkine;
KW bonzo; ss.
XX
OS Homo sapiens.

XX Key Location/Qualifiers
FT 15..779
FT /*tag= a
FT /product= "SExCkine"
FT /transl_except= "(pos: 321..323, aa: Glu)"
FT /transl_except= "(pos: 381..383, aa: Thr)"
FT sig_peptide 15..101
FT /*tag= c
XX
PN WO200138874-A1.

XX 31-MAY-2001.
PD
XX 22-NOV-2000; 2000WO-US32095.
PF
XX 24-NOV-1999; 99US-0448725.
PR
XX (MILL-) MILLENNIUM PHARM INC.

XX Briskin MJ;
PI
XX WPI; 2001-367732/38.
DR
DR P-PSDB; AAB84199.

XX Isolating genes encoding G-protein-coupled receptor ligand, comprises
PT expressing an exogenous nucleic acid library in a primary cell pool,
PT selecting and growing cells that encode ligand, and combining them into
PT a secondary cell pool
XX
PS Example; Fig 2; 43pp; English.

XX

CC The specification describes a method for isolating nucleic acid
CC encoding a G protein-coupled receptor (GPCR) ligand. The method comprises
CC expressing an expression library having exogenous nucleic acids
CC inserted into a primary pool of prokaryotic cells, to isolate a pool
CC of proteins comprising the ligand, selecting and growing cells that
CC encode the pool of proteins to produce individual colonies, and
CC combining them into a secondary pool of cells for recovering the
CC nucleic acid. The method is useful for isolating a nucleic acid,
CC preferably a cDNA encoding a ligand for a GPCR, such as a chemokine
CC receptor. The present sequence encodes a spleen extracted chemokine
CC (SExCkine), a ligand for bonzo. SExCkine was isolated using the method
CC of the invention.

XX Sequence 1763 BP; 374 A; 514 C; 395 G; 480 T; 0 other;

Query Match 2.7%; Score 28; DB 22; Length 1763;
Best Local Similarity 100.0%; Pred. No. 0.00046;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 ccatgttgaccaggctggtctcgaactc 696
|||||

Fri May 3 11:02:10 2002

us-09-671-050-9.oli20.rng

Db 1256 ccatgttgaccaggctggtctcgaactc 1283

RESULT 37

AAF85365

ID AAF85365 standard; cDNA; 1763 BP.

XX

AC AAF85365;

XX

XX 23-JUL-2001 (first entry)

XX

DE Nucleotide sequence of a human spleen extracted chemokine (SExCkine).

XX

XX Bonzo; CXc chemokine receptor; inflammatory disease; cancer; infection;

KW SExCkine; spleen extracted chemokine; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 15..779

FT /*tag= a

FT /product= "spleen extracted chemokine"

FT sig_peptide 15..95

FT /*tag= b

XX

PN WO200137872-A1.

XX

XX 31-MAY-2001.

PD

XX

XX 22-NOV-2000; 2000WO-US32206.

PF

XX

XX 24-NOV-1999; 99US-0449437.

PR

XX

PA (MILL-) MILLENNIUM PHARM INC.

XX

XX Briskin MJ, Murphy KE, Wilbanks AM, Wu L;

PI

XX

XX WPI; 2001-343947/36.

DR

XX

XX P-PSDB; AAB68423.

DR

XX

XX Identifying agents (especially antibodies) which bind to the CXc

PT chemokine receptor Bonzo, and which may be used to treat e.g. cancers

PT and inflammation -

PT

XX

PS Example 1; Fig 3; 191pp; English.

PS

XX

XX The present sequence encodes a human spleen extracted chemokine

CC (SExCkine). SExCkine is a ligand for Bonzo polypeptides. Bonzo is a

CC CXc chemokine receptor. The specification describes a method for

CC identifying agents (especially antibodies) which bind to Bonzo and

CC inhibit the binding of a ligand (especially SExCkine) and the agents

CC per se. The agents identified may be used for the treatment of a

CC disorder/disease related to aberrant Bonzo expression and activity,

CC such as inflammatory disease, cancers and/or infections (e.g. viral,

CC bacterial and fungal infections).

XX

SQ Sequence 1763 BP; 374 A; 514 C; 395 G; 480 T; 0 other;

Query Match 2.7%; Score 28; DB 22; Length 1763;

Best Local Similarity 100.0%; Pred. No. 0.00046;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 ccatgttgaccaggctggtctcgaactc 696

Db 1256 ccatgttgaccaggctggtctcgaactc 1283

RESULT 38

AAH18124/C

ID AAH18124 standard; cDNA; 1791 BP.

XX

AC AAH18124;

XX

DT 26-JUN-2001 (first entry)

XX

DE Human cDNA sequence SEQ ID NO:17990.

XX

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX

OS Homo sapiens.

XX

PN EP1074617-A2.

XX

PD 07-FEB-2001.

XX

XX 28-JUL-2000; 2000EP-0116126.

PF

XX

XX 29-JUL-1999; 99JP-0248036.

PR

XX 27-AUG-1999; 99JP-0300253.

PR

XX 11-JAN-2000; 2000JP-0118776.

PR

XX 02-MAY-2000; 2000JP-0183767.

PR

XX 09-JUN-2000; 2000JP-0241899.

PR

XX

PA (HELI-) HELIX RES INST.

XX

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

PI

XX

XX WPI; 2001-318749/34.

DR

XX

XX Primer sets for synthesizing polynucleotides, particularly the 5602

PT full-length cDNAs defined in the specification, and for the detection

PT and/or diagnosis of the abnormality of the proteins encoded by the

PT full-length cDNAs -

XX

XX Claim 8; SEQ ID 17990; 2537pp + CD ROM; English.

PS

XX

XX The present invention describes primer sets for synthesising 5602

CC full-length cDNAs defined in the specification. Where a primer set

CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary

CC to the complementary strand of a polynucleotide which comprises one of

CC the 5602 nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination

CC of an oligonucleotide comprising a sequence complementary to the

CC complementary strand of a polynucleotide which comprises a 5'-end

CC sequence and an oligonucleotide comprising a sequence complementary to a

CC polynucleotide which comprises a 3'-end sequence, where the

CC oligonucleotide comprises at least 15 nucleotides and the combination of

CC the 5'-end sequence/3'-end sequence is selected from those defined in

CC the specification. The primer sets can be used in antisense therapy and

CC in gene therapy. The primers are useful for synthesising polynucleotides,

CC particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by

CC the full-length cDNAs. The primers allow obtaining of the full-length

CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to

CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632

CC represent oligonucleotides, all of which are used in the exemplification

CC of the present invention.

XX

SQ Sequence 1791 BP; 484 A; 402 C; 433 G; 472 T; 0 other;

Query Match 2.7%; Score 28; DB 22; Length 1791;

Best Local Similarity 100.0%; Pred. No. 0.00046;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 ccatgttgaccaggctggtctcgaactc 696

Db 1767 CCATGTTGACCAGGCTGGTCTCGAACTC 1740

RESULT 39

AAAC59834

ID AAC59834 standard; DNA; 1792 BP.

XX AAC59834;
XX
XX
DT 26-JAN-2001 (first entry)
XX
DE Human secreted protein encoding DNA clone vq19 1.
XX
XX
KW Secreted protein; human; autoimmune disorder; multiple sclerosis; ulcer;
KW systemic lupus erythematosus; rheumatoid arthritis; anaemia; stroke;
KW haematopoiesis regulation; tissue regrowth; wound healing; haemophilia;
KW Alzheimer's disease; Parkinson's disease; Shy-drager syndrome; cancer;
KW contraceptive; infection; growth inhibition; hyperproliferative disorder;
KW psoriasis; ds.
XX
OS Homo sapiens.
XX
XX WO200055375-A1.
PN
XX
PD 21-SEP-2000.
XX
XX 17-MAR-2000; 2000WO-US07285.
PF
XX
PR 17-MAR-1999; 99US-0124808.
PR 17-MAR-1999; 99US-0124916.
PR 17-AUG-1999; 99US-0149639.
PR 01-OCT-1999; 99US-0157247.
PR 29-NOV-1999; 99US-0167824.
PR 15-FEB-2000; 2000US-0182711.
XX
XX (ALPH-) ALPHAGENE INC.
PA
XX
XX Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;
PI
XX
XX WPI; 2000-638211/61.
DR
DR P-PSDB; AAB34733.
XX
XX
XX Novel proteins and polypeptides useful for the treatment of e.g
PT multiple sclerosis, systemic lupus erythmatosus, rheumatoid arthritis,
PT cancer, Alzheimer's disease, Parkinson's disease, stroke, anemia and
PT ulcers -
PS
XX Claim 102; Page 446-447; 493pp; English.

This invention relates to 59 human secreted proteins and the nucleotide sequences encoding them. Sequences AAC59788-C59846 and AAB34687-B34745 represent the proteins and their encoding nucleotide sequences, and sequences AAB34746-B34771 represent fragments of the proteins. Probes for the DNA sequences are represented by sequences AAC59847-C59596. The proteins exhibit neuroprotective, dermatological, immunosuppressive, antiinflammatory, antianaemic, nootropic, antiparkinsonian, cerebroprotective, haemostatic, vulnerable, cytostatic, antipsoriatic, antibacterial, virucide, and fungicide activity. The proteins and nucleotide sequences are useful as nutritional sources or supplements and in research. The proteins are useful for treating immune deficiency and disorders, which may be genetic or resulting from infections, autoimmune disorders such as multiple sclerosis, systemic lupus erythmatosus, rheumatoid arthritis, and for treating myeloid or lymphoid cell deficiencies such as anaemias by regulating haematopoiesis. The proteins are also useful in compositions for bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, for wound healing, tissue repair and replacement and in the treatment of wounds, incisions and ulcers. Other uses include in the treatment of central and peripheral nervous system and neuropathies such as Alzheimer's and Parkinson's diseases and Shy-Drager syndrome, and mechanical and traumatic disorders, such as spinal cord disorders, head trauma and stroke. The proteins may also be used as a contraceptive, and for treating coagulation disorders such as haemophilias. The protein and nucleotide sequences with cadherin activity are useful for treating cancer. Other uses for the protein include for inhibiting the growth, infection or function of, or killing, infectious agents such as bacteria, virus, fungi and other parasites, for effecting bodily characteristics such as height, weight, hair colour, effecting biorhythms or cardiac cycles or rhythms, effecting metabolism, catabolism, anabolism,

CC processing, utilization, storage or elimination of dietary fat, lipid,
CC protein, carbohydrate, vitamins, minerals, cofactors, effecting
CC behavioural characteristics, providing analgesic effects and for treating
CC hyperproliferative disorders such as psoriasis.
XX
SQ Sequence 1792 BP; 417 A; 507 C; 387 G; 481 T; 0 other;

Query Match 2.7%; Score 28; DB 21; Length 1792;
Best Local Similarity 100.0%; Pred. No. 0.00046;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 ccattgttgaccaggctggtctcgaaactc 696
||||||||||||||||||||||||||
Db 1241 ccattgttgaccaggctggtctcgaaactc 1268

RESULT 40
AAH18088
ID AAH18088 standard; cDNA; 1850 BP.
XX
AC AAH18088;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:17938.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs -
PS Claim 8; SEQ ID 17938; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by

CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX Sequence 1850 BP; 473 A; 444 C; 390 G; 543 T; 0 other;

Query Match 2.7%; Score 28; DB 22; Length 1850;
Best Local Similarity 100.0%; Pred. No. 0.00046;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 ccatgttgaccaggctggtctcgaactc 696
|||||
Db 1266 ccatgttgaccaggctggtctcgaactc 1293

RESULT 41
AAC76936
ID AAC76936 standard; cDNA; 1902 BP.
XX
AC AAC76936;

DT 08-FEB-2001 (first entry)
XX Human ORFX ORF2491 polynucleotide sequence SEQ ID NO:4981.
DE

XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive; ss.

XX Homo sapiens.
OS
XX
PN WO200058473-A2.

XX
PD 05-OCT-2000.
XX
PF 31-MAR-2000; 2000WO-US08621.
XX
PR 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.

XX (CURA-) CURAGEN CORP.
PA
XX
XX Shimkets RA, Leach M;
PI
XX
DR WPI; 2000-602362/57.
DR P-PSDB; AAB42727.

XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
PT

XX Claim 5; Page 4155-4157; 5507pp; English.
PS
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC

CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antithyroid; and antianaemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ Sequence 1902 BP; 366 A; 541 C; 446 G; 548 T; 1 other;

Query Match 2.7%; Score 28; DB 21; Length 1902;
Best Local Similarity 100.0%; Pred. No. 0.00046;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 ccatgttgaccaggctggtctcgaactc 696
|||||
Db 652 ccatgttgaccaggctggtctcgaactc 679

RESULT 42

AAH17566
ID AAH17566 standard; cDNA; 2252 BP.

XX AAH17566;
AC
XX 26-JUN-2001 (first entry)
DT
XX Human cDNA sequence SEQ ID NO:17050.
DE

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
KW
XX Homo sapiens.
OS
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.
PA
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX

PS Claim 8; SEQ ID 17050; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesising 5602
CC full-length cDNAs defined in the specification. Where a primer set

comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

Sequence 2252 BP; 435 A; 692 C; 571 G; 554 T; 0 other;

Query Match 2.7%; Score 28; DB 22; Length 2252;
Best Local Similarity 100.0%; Pred. No. 0.00046;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 ccattgtgaccaggctggtctcgaaactc 696
|||||
Db 1747 ccattgtgaccaggctggtctcgaaactc 1774

RESULT 43
AAZ25332
ID AAZ25332 standard; cDNA; 2309 BP.
AC AAZ25332;
DT 20-DEC-1999 (first entry)
DE Human chemokine alpha-5 nucleotide sequence.
KW Human; chemokine alpha-5; CKalpha-5; diagnosis; detection; immunity;
KW immune system related disorder; solid tumour; angiogenesis; lymphoma;
KW infection; autoimmune disease; lymphocytic leukaemias; wound healing;
KW fibrotic disorder; liver cirrhosis; osteoarthritis; pulmonary fibrosis;
KW haematopoiesis; regulation; autoimmunity; arthritis; leukaemias;
KW immunosuppression; inflammatory bowel disease; myelo suppression; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 542..1306
FT /*tag= a
FT sig_peptide 542..622
FT /*tag= b
FT mat_peptide 623..1303
FT /*tag= c
XX
PN ~ WO9927078-A1.
XX
PD 03-JUN-1999.
XX
PF 18-NOV-1998; 98WO-US24619.
XX
PR 21-NOV-1997; 97US-0066369.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Wei Y, Ni J, Li Y, Rosen CA, Ruben SM;
XX WPI; 1999-590675/50.
DR

P-PSDB; AAY42165.
New isolated chemokine alpha-5
Claim 2; Fig 1; 86pp; English.
The present sequence encodes human chemokine alpha-5 (CKalpha-5). The novel CKalpha-5 polypeptides have chemokine activity. It can be used to treat solid tumours, angiogenesis, chronic and acute infections, e.g. myobacterial infections, autoimmune diseases and lymphocytic leukaemias, stimulate wound healing, fibrotic disorders, (e.g. liver cirrhosis, osteoarthritis and pulmonary fibrosis), killing the larvae of parasites that invade tissues (as in schistosomiasis, trichinosis and ascariasis), regulate haematopoiesis, re-vascularisation of damaged limbs (from injury or disease), dysregulation of immune cell function (e.g. autoimmunity, arthritis, leukaemias, lymphomas, immunosuppression, immunity, humoral immunity, inflammatory bowel disease, or myelo suppression). Antagonists to CKalpha-5 can be used in the treatment of autoimmune and chronic inflammatory and infective diseases (e.g. multiple sclerosis and insulin-dependent diabetes), infectious diseases (e.g. silicosis, sarcoidosis, and idiopathic pulmonary fibrosis), idiopathic hyper-eosinophilic syndrome, endotoxic shock, atherosclerosis, histamine-mediated allergic reactions and immunological disorders (e.g. late phase allergic reactions, chronic urticaria, and atopic dermatitis), IgE-mediated allergic reactions (e.g. allergic asthma, rhinitis and eczema), to treat chronic and acute inflammation, rheumatoid arthritis, bone marrow failure, (e.g. aplastic anaemia and myelodysplastic syndrome), and subepithelial basement membrane fibrosis. The products can also be used for detection, diagnosis and drug screening.

Sequence 2309 BP; 462 A; 706 C; 581 G; 560 T; 0 other;

Query Match 2.7%; Score 28; DB 20; Length 2309;
Best Local Similarity 100.0%; Pred. No. 0.00046;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 ccattgtgaccaggctggtctcgaaactc 696
|||||
Db 1782 ccattgtgaccaggctggtctcgaaactc 1809

RESULT 44
AAZ71606
ID AAA71606 standard; DNA; 2575 BP.
XX
AC AAA71606;
XX
DT 11-DEC-2000 (first entry)
XX
DE Human CXC family chemokine-like protein encoding DNA SEQ ID NO: 1.
XX
KW Human; CXC family; chemokine-like; disease; treatment; CL protein; ds.
XX
OS Homo sapiens.
XX
PN WO200044784-A1.
XX
PD 03-AUG-2000.
XX
PF 21-JAN-2000; 2000WO-IL00043.
XX
PR 27-JAN-1999; 99IL-0128252.
XX
PA (COMP-) COMPUGEN LTD.
XX
PI Mintz L, Savitzky K;
XX
DR WPI; 2000-491238/43.
DR P-PSDB; AAB10491.
XX
PT New human chemokine like (CL) nucleic acids and proteins, useful for

PT diagnosing and treating diseases resulting from inherited defective CL
PT sequences -
XX
PS Claim 1(i); Page 47-48; 55pp; English.
XX
CC This invention describes novel human chemokine like (CL) nucleic acids
CC (N1), their complementary sequences (N2) and encoded amino acids (P1).
CC An expression vector or P1 is useful for treating diseases which
CC can be ameliorated or cured by raising the level of the CL protein. N2,
CC the expression vector or an antibody raised against P1 is useful for
CC treating diseases which can be ameliorated or cured by decreasing the
CC level of the CL protein. The CL nucleic acids may be used to detect and
CC quantitate expression of CL protein in biopsied tissues. The CL nucleic
CC acids may also be used in methods for diagnosing diseases resulting from
CC inherited defective CL sequences. The antibodies are also useful for
CC detecting the CL protein in biological samples. This sequence encodes a
CC human CXC family chemokine-like protein which is described in the method
CC of the invention.
XX
SQ Sequence 2575 BP; 517 A; 744 C; 629 G; 685 T; 0 other;

Query Match 2.7%; Score 28; DB 21; Length 2575;
Best Local Similarity 100.0%; Pred. No. 0.00046;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 ccattgtgaccaggctggtctcgaaactc 696
|||||
Db 1340 ccattgtgaccaggctggtctcgaaactc 1367
|||||

RESULT 45
AAH18187
ID AAH18187 standard; cDNA; 2797 BP.
XX
AC AAH18187;
XX
DT 26-JUN-2001 (first entry)
XX
XX Human cDNA sequence SEQ ID NO:18096.
DE
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
KW
XX Homo sapiens.
OS
XX EP1074617-A2.
PN
XX 07-FEB-2001.
PD
XX 28-JUL-2000; 2000EP-0116126.
PF
XX 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
PA
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI; 2001-318749/34.
DR
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 8; SEQ ID 18096; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesising 5602
CC full-length cDNAs defined in the specification. Where a primer set

CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 2797 BP; 621 A; 721 C; 655 G; 800 T; 0 other;

Query Match 2.7%; Score 28; DB 22; Length 2797;
Best Local Similarity 100.0%; Pred. No. 0.00046;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 ccattgtgaccaggctggtctcgaaactc 696
|||||
Db 1529 ccattgtgaccaggctggtctcgaaactc 1556
|||||

RESULT 46
AAH71607
ID AAH71607 standard; DNA; 2799 BP.
XX
AC AAH71607;
XX
DT 11-DEC-2000 (first entry)
XX
XX Human CXC family chemokine-like protein encoding DNA SEQ ID NO: 2.
DE
XX Human; CXC family; chemokine-like; disease; treatment; CL protein; ds.
KW
XX Homo sapiens.
OS
XX WO200004784-A1.
PN
XX 03-AUG-2000.
PD
XX 21-JAN-2000; 2000WO-IL00043.
PF
XX 27-JAN-1999; 99IL-0128252.
PR
XX (COMP-) COMPUGEN LTD.
PA
XX Mintz L, Savitzky K;
PI
XX WPI; 2000-491238/43.
DR
XX P-PSDB; AAB10492.
XX
XX New human chemokine like (CL) nucleic acids and proteins, useful for
PT diagnosing and treating diseases resulting from inherited defective CL
PT sequences -
XX
PS Claim 1(i); Page 48-49; 55pp; English.
XX
CC This invention describes novel human chemokine like (CL) nucleic acids
CC (N1), their complementary sequences (N2) and encoded amino acids (P1).
CC An expression vector or P1 is useful for treating diseases which
CC can be ameliorated or cured by raising the level of the CL protein. N2,
CC the expression vector or an antibody raised against P1 is useful for
CC diagnosing and treating diseases resulting from inherited defective CL
CC sequences -

CC treating diseases which can be ameliorated or cured by decreasing the
CC level of the CL protein. The CL nucleic acids may be used to detect and
CC quantitate expression of CL protein in biopsied tissues. The CL nucleic
CC acids may also be used in methods for diagnosing diseases resulting from
CC inherited defective CL sequences. The antibodies are also useful for
CC detecting the CL protein in biological samples. This sequence encodes a
CC human CXC family chemokine-like protein which is described in the method
CC of the invention.
XX
SQ Sequence 2799 BP; 539 A; 825 C; 707 G; 728 T; 0 other;

Query Match 2.7%; Score 28; DB 21; Length 2799;
Best Local Similarity 100.0%; Pred. No. 0.00046;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 ccattgtgaccaggctggtctcgaactc 696
|||||
Db 1564 ccattgtgaccaggctggtctcgaactc 1591

RESULT 47
AAH17249
ID AAH17249 standard; cDNA; 2845 BP.
XX
AC AAH17249;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:16639.
XX

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX Homo sapiens.

OS
XX EP1074617-A2,
PN 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX

PA (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI; 2001-318749/34.

XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -

PS Claim 8; SEQ ID 16639; 2537pp + CD ROM; English.

XX
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in

CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX

SQ Sequence 2845 BP; 839 A; 628 C; 661 G; 717 T; 0 other;

Query Match 2.7%; Score 28; DB 22; Length 2845;
Best Local Similarity 100.0%; Pred. No. 0.00046;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 663 gtttcgccatgttgaccaggctggtctc 690
|||||
Db 195 gtttcgccatgttgaccaggctggtctc 222

RESULT 48
AAF21279/C
ID AAF21279 standard; DNA; 3585 BP.

XX
AC AAF21279;

XX
DT 14-MAR-2001 (first entry)

XX
DE Human low adenosine antisense oligonucleotide related sequence #2846.

XX
KW Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
KW human; airway disorder; bronchoconstriction; lung inflammation;
KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;
KW immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
KW respiratory obstruction; pulmonary obstruction; impeded respiration;
KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;
KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
cancer; ss.

XX Homo sapiens.

XX WO200062736-A2.

XX 26-OCT-2000.

XX 24-MAR-2000; 2000WO-US08020.

XX 06-APR-1999; 99US-0127958.

XX (UYEC-) UNIV EAST CAROLINA.

XX (NYCE/) NYCE J W.

XX Nyce JW;

XX WPI; 2000-679539/66.

XX Low adenosine (A) content antisense oligonucleotides which do not
PT trigger adenosine receptors during metabolism, useful e.g. for treating
PT cancers and respiratory obstructions -

PS Disclosure; Page 1258-1259; 1592pp; English.

XX The present invention describes low adenosine (A) content antisense
CC oligonucleotides and compositions (I) comprising them. In the antisense
CC oligonucleotides the A is replaced by a 'Universal' or alternative base.
CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
CC immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.
CC The antisense oligonucleotides and (I) can be used to down-regulate the

CC expression and or activity of target polypeptides associated with
CC lung/respiratory disorders and malignancies, such as stimulating and
CC activating peptide factors and transmitters, transcription factors,
CC immunoglobulins and antibodies, antibody receptors, cytokines and
CC chemokines, endogenously produced specific and non-specific enzymes,
CC binding proteins, adhesion molecules and their receptors, cytokine and
CC chemokine receptors, adenosine receptors, bradykinin receptors, central
CC nervous system (CNS) and peripheral nervous and non-nervous system
CC receptors, CNS and peripheral nervous and non-nervous system peptide
CC transmitters, defensins, growth factors, vasoactive peptides and
CC receptors, binding proteins and malignancy associated proteins. The
CC antisense oligonucleotides may be used in this way to treat disorders
CC including respiratory obstruction (especially pulmonary obstruction
CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies)
CC and/or surfactant hypoproduction which are associated with a disease or
CC condition selected from pulmonary vasoconstriction, inflammation,
CC allergies, asthma, impeded respiration, respiratory distress syndrome
CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
CC fragments and antisense oligonucleotides used in the exemplification of
CC the present invention.

XX
SQ Sequence 3585 BP; 864 A; 945 C; 1066 G; 710 T; 0 other;

Query Match 2.7%; Score 28; DB 21; Length 3585;
Best Local Similarity 100.0%; Pred. No. 0.00046;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 ccattgttgaccaggctgtctcgaaactc 696
|||||
Db 1506 CCATGTTGACGAGGCTGGTCTCGAACTC 1479

RESULT 49
AAA35157/c
ID AAA35157 standard; DNA; 3585 BP.
XX
AC AAA35157;
XX
28-JUL-2000 (first entry)
XX
DE Human adenosine receptor related polynucleotide 2nd SEQ ID NO:31.
XX
XX Human; adenosine receptor; low adenosine antisense oligonucleotide;
KW phosphorothioate; impaired respiration; inflammation; allergy;
KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
KW antiallergic; antiasthmatic; cytotstatic; analgesic; impaired airway;
KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;
KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.
XX
OS Homo sapiens.
XX
XX WO200009525-A2.
PN
XX
XX 24-FEB-2000.
PD
XX
XX 03-AUG-1999; 99WO-US17712.
PF
XX
XX 03-AUG-1998; 98US-0095212.
PR
XX
XX (UYEC-) UNIV EAST CAROLINA.
PA
XX
XX Nyce JW;
PI
XX
XX WPI; 2000-205971/18.
DR
XX
XX New antisense oligonucleotides useful for treating e.g. pulmonary
PT vasoconstriction, inflammation, allergies, asthma, hypertension,
PT

PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
PT cancers
XX
XX Disclosure; Page 1175-1176; 1343pp; English.
XX
CC The present invention describes a new composition comprising an
CC antisense oligonucleotide (ON) with low adenosine (up to 15%), which
CC targets nucleic acids involved in bronchoconstriction, allergies, and/or
CC inflammation. The ON can have antiinflammatory, antiallergic,
CC antiasthmatic, cytotstatic and analgesic activities. The compositions are
CC useful for the treatment of diseases associated with inflammation,
CC impaired airways, including lung disease and diseases whose secondary
CC effects afflict the lungs of a subject. They can be used for treating
CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,
CC asthma, impeded respiration, respiratory distress syndrome, pain, cystic
CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive
CC pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
CC carcinomas, and cancers which may metastasise to the lungs, including
CC breast and prostate cancer. The reduction of the adenosine content of
CC the ONs reduces side effects. The A-containing ONs break down with the
CC release of deoxyadenosine which activates adenosine receptors causing the
CC bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the
CC nucleotide sequences given in the sequence listing from the present
CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last
CC 185 sequences are also called SEQ ID NO:1 to 185, but the sequences
CC differ from the previously named sequences. SEQ ID NO:11 to 1680
CC (AAA32323 to AAA33992) are specifically claimed ONs from the present
CC invention. N.B. Sequences given in the disclosure of the present
CC invention do not match up with their corresponding SEQ ID NO: sequences
CC given in the sequence listing.

XX
SQ Sequence 3585 BP; 864 A; 945 C; 1066 G; 710 T; 0 other;

Query Match 2.7%; Score 28; DB 21; Length 3585;
Best Local Similarity 100.0%; Pred. No. 0.00046;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 ccattgttgaccaggctgtctcgaaactc 696
|||||
Db 1506 CCATGTTGACGAGGCTGGTCTCGAACTC 1479

RESULT 50
AAH23153/c
ID AAH23153 standard; DNA; 3586 BP.
XX
AC AAH23153;
XX
XX 17-SEP-2001 (first entry)
DT
XX Nitric oxide synthase gene partial sequence.
DE
XX Hypertension; renal disease; non insulin dependent diabetes mellitus;
KW polymorphism; nitric oxide synthase; breast; lung; prostate; cancer;
KW nephrotropic; vasodilatory; renal injury; NIDDM; SNP; NOS; ds.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH promoter 1465..3472
FT /*tag= a
FT CDS 3494..3586
FT /*tag= b
FT /*note= "partial coding sequence"
XX
XX WO200153537-A2.
PN
XX
XX 26-JUL-2001.
PD
XX
XX 24-JAN-2001; 2001WO-US02260.
PF
XX
XX 24-JAN-2000; 2000US-0177775.
PR

25-JUL-2000; 2000US-0220662.
(DZGE-) DZGENES LLC.
Moskowitz DW;
WPI; 2001-451916/48.
P-PSDB; AAB85401.
The diagnosis of disorders, i.e. hypertension and non insulin dependent diabetes mellitus comprises the identification of single polymorphism(s) in a nitric oxide synthase gene -
Claim 1; Page 104-106; 119pp; English.
The invention relates to diagnosing a genetic predisposition to disease, condition or disorder i.e. hypertension, non insulin dependent diabetes mellitus (NIDDM), end stage renal disease due to hypertension and/or NIDDM mellitus that involves detecting the presence or absence of single nucleotide polymorphisms (SNP) in the nitric oxide synthase gene (NOS). The method is also useful for the treatment or prophylaxis of breast, lung or prostate cancer. The method allows the identification of people with a predisposition to disorders, before the disease progresses to an end stage. The present sequence represents a partial sequence of the human nitric oxide synthase gene.
Sequence 3586 BP; 864 A; 945 C; 1066 G; 710 T; 1 other;

Query Match	2.7%;	Score 28;	DB 22;	Length 3586;
Best Local Similarity	100.0%;	Pred. No. 0.00046;		
Matches 28;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy 669	ccatgttgaccaggctgggtctcgaactc	696		
Db 1506	CCATGTTGACCAGGCTGGTCTCGAACTC	1479		

Search completed: May 2, 2002, 19:45:11
Job time: 3700 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 13:50:18 ; Search time 2919.35 Seconds
(without alignments)
5882.667 Million cell updates/sec

Title: US-09-671-050-9
Perfect score: 1041
Sequence: 1 atggaaagtatgaaaaatt.....aggtacttcgcgtcaaaagt 1041

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues 2944280

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rod:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rod:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

8

Result No.	Score	Query Match	Length	DB	ID	Description
1	1041	100.0	1041	6	AX107720	AX107720 Sequence
2	1025	98.5	1068	6	AX107714	AX107714 Sequence
3	839	80.6	945	6	AX107722	AX107722 Sequence
4	823	79.1	972	6	AX107716	AX107716 Sequence
5	757	72.7	1083	6	AX166534	AX166534 Sequence
6	709	68.1	1819	6	AX107724	AX107724 Sequence
7	537	51.6	911	6	AX056404	AX056404 Sequence
8	454.4	43.7	561	6	AX107712	AX107712 Sequence
9	454.4	43.7	594	6	AX107718	AX107718 Sequence
10	453.2	43.5	2615	6	AX056405	AX056405 Sequence
11	343.8	33.0	1363	9	HSSTHPKB	X66358 H.sapiens m
12	251.8	24.2	3153	4	AB029045	AB029045 Oryctolag
13	243.8	23.4	1518	10	AB029067	AB029067 Mus muscu
14	243.8	23.4	3988	10	AB029066	AB029066 Mus muscu
15	243.8	23.4	4022	10	AB029065	AB029065 Mus muscu
16	241.2	23.2	1993	9	HSU35146	U35146 Human p56 K
17	202.6	19.5	1513	9	AF130372	AF130372 Homo sapi
18	202.6	19.5	1773	6	AX166533	AX166533 Sequence
19	201	19.3	1974	3	AF326965	AF326965 Trypanoso
20	199	19.1	80362	9	AC079615	AC079615 Homo sapi
21	199	19.1	155028	33	AC037489	Ac037489 Homo sapi
22	199	19.1	164281	2	AC092672	AC092672 Homo sapi
23	191	18.3	1732	10	AF112183	AF112183 Rattus no
24	191	18.3	1738	10	AF112184	AF112184 Rattus no
c 25	186.6	17.9	101340	2	AC018104	AC018104 Drosophil
c 26	186.6	17.9	190668	3	AC007419	AC007419 Drosophil
27	186.6	17.9	259973	3	AE003611	AE003611 Drosophil
28	177.2	17.0	3399	9	HSTXP3	Y15057 Homo sapien
29	175.2	16.8	1866	6	AX056407	AX056407 Sequence
30	166	15.9	960	3	TCU69958	U69958 Trypanosoma
31	163.8	15.7	903	8	AF026545	AF026545 Pneumocys
32	162	15.6	1476	6	AR087351	AR087351 Sequence
33	162	15.6	1476	9	HSCDK2MR	X61622 H.sapiens C
34	162	15.6	1635	6	A84497	A84497 Sequence 11
35	162	15.6	1635	6	A84499	A84499 Sequence 11
36	161.6	15.5	1297	9	HSCDK2	X62071 H.sapiens c
37	161	15.5	921	3	DDICRP	L00652 Dictyosteli
38	160.4	15.4	2213	9	HUMCDG2A	M68520 Human cdc2-
39	160.4	15.4	2264	9	BC003065	BC003065 Homo sapi
40	155.6	14.9	2213	10	CGR223949	AJ223949 Cricetulu
41	154.6	14.9	885	3	AF076534	AF076534 Cryptospo
42	154.2	14.8	1050	9	HSCDC2	X05360 Human CDC2
43	154.2	14.8	1050	9	HSCDC2R	Y00272 Human cell
44	153	14.7	3321	3	LME293284	AJ293284 Leishmani
45	152.4	14.6	1170	5	AF159158	AF159158 Rana dybo

ALIGNMENTS

RESULT 1	AX107720	AX107720	1041 bp	DNA	PAT	30-APR-2001
AX107720	Sequence 9	from Patent	WO0123579.			
LOCUS	AX107720					
DEFINITION	AX107720					
ACCESSION	AX107720.1	GI:13923201				
VERSION						
KEYWORDS	human.					
SOURCE	Homo sapiens					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1 (bases 1 to 1041)					
AUTHORS	Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.					
TITLE	Human kinase proteins and polynucleotides encoding the same					
JOURNAL	Patent: WO 0123579-A 9 05-APR-2001;					
FEATURES	Lexicon Genetics Incorporated (US)					
source	Location/Qualifiers					
	1..1041					
	/organism="Homo sapiens"					


```
QY 541 tcagtcgatatatgggctattggtgtgtgttttgcagagtcctgcagggccagccactg 600
|||||
Db 541 TCAGTCGATATATGGGCTATTGGTGTGTGTTTTGCAGAGCTCCTGACAGGCCAGCCACTG 600

QY 601 tggcctggaaaaatcagatgtggaccaactttatctgataatcagaaacactagtagagacg 660
|||||
Db 601 TGCCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGTAGAGACG 660

QY 661 gggtttcgccaattgtgaccaggctggtctgcgaactcttgacgtcaagtgcaccctgcc 720
|||||
Db 661 GGGTTTCGCCCATGTTGACCAGGCTGGTCTCGAACTCTTGACGTCAAGTGATCCACCTGCC 720

QY 721 gtgacctctcaaaagtctgggaattacaggaaaaattaatcccaagacatcaatcaacttt 780
|||||
Db 721 GTAGCCTCTCAAAGTCTGGAATTACAGGAAAAATTAATCCCAAGACATCAATCAATCTTT 780

QY 781 aaaagtaacgggtttttccatggcatcagtatatacctgagccagaagacatggaaccttt 840
|||||
Db 781 AAAAGTAAACGGGTTTTCCATGGCATCAGTATACCTGAGCCAGAACATGGAAACTCTT 840

QY 841 gaggaanaagttctcagatgttcatcctgtggtctggaactcctggagagctcctactttgaa 900
|||||
Db 841 GAGGAAAAGTTCTCAGATGTTTCATCCTGTGGCTCTGAACTTTCATGGAAGGGTGTCTGAAG 900

QY 901 atgaatccagatgacagattaacctgttcccactcctggagagctcctactttgattct 960
|||||
Db 901 ATGAATCCAGATGACAGATTAACTGTTCCTCCAACTCCTGGAGAGCTCCTACTTTGATTCT 960

QY 961 tttcaagaggcccaaatataaagaaaaagcacgttaataaggaagaaacagagacgcca 1020
|||||
Db 961 TTTCAAGAGGCCCAAAATTAAGAGAAAAGCACGTAATGAAGGAAGAAACAGAGACGCCAA 1020

QY 1021 caggt 1025
|||||
Db 1021 CAGGT 1025

RESULT 3
AX107722 945 bp DNA PAT 30-APR-2001
LOCUS
DEFINITION Sequence 11 from Patent WO0123579.
ACCESSION AX107722
VERSION AX107722.1 GI:13923202
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 945)
AUTHORS Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and
Sands,A.T.
TITLE Human kinase proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0123579-A 11 05-APR-2001;
Lexicon Genetics Incorporated (US)
FEATURES
source
1..945
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 323 a 165 c 202 g 255 t
ORIGIN
Query Match 80.6%; Score 839; DB 6; Length 945;
Best Local Similarity 90.8%; Pred. No. 3.2e-187;
Matches 945; Conservative 0; Mismatches 0; Indels 96; Gaps 1;

QY 1 atggaaaaagtatgaaaaattagctaaactggagagggtcttatggggtgtattcaaa 60
|||||
Db 1 ATGGAAAAAGTATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTGTATTTCAAA 60

QY 61 tgcagaaacaaacctctggacaagttagctgttaaaaaattgtggaatctgaagat 120
|||||
```

```
Db 61 TGCAGAAACAAAACCTCTGGACAAAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGRAGAT 120

QY 121 gatcctgttgttaagaaaaatagcactaagagaaatacgtatgttgaagcaattaaaaacat 180
|||||
Db 121 GATCCTGTTGTTAAGAAAAATAGCACCTAAGAGAAATACGTATGTGGAAGCAATTAAACAT 180

QY 181 ccaaatcttglgaacctcatcgcaggtgttccaggagaaaaaaggaaaaatgcatttagtttt 240
|||||
Db 181 CCAAAATCTTGTGAACCTCATCGAGGTGTTCCAGGAGAAAAAGGAAAAATGCATTTAGTTT 240

QY 241 gaatactgtgatcatcacacttttaaatgagctggaaagaaaccccaaatggagttgctgat 300
|||||
Db 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGAT 300

QY 301 ggagtgcataaaagcgtattatggcaaacacttcaagctcttaatttctgcataatacat 360
|||||
Db 301 GGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTCATATACAT 360

QY 361 aactgtattcacagagatataaaacctgaaaaatattcttaataactaagcaaggaataatc 420
|||||
Db 361 AACTGTATTACAGAGATATAAAACCTGAAAAATATTCTAATAACTAAGCAAGGAATAATC 420

QY 421 aagatttgtgacttcgggtttgcacaaattctgattccaggagatgcctacacgattat 480
|||||
Db 421 AAGATTGTGACTTCGGGTTTGCACAAATCTGATTCAGGAGATGCCTACACCGATTAT 480

QY 481 gtagtcacgagatgtgaccagctcctgaaacttcttgtgggagagatactcagtatggtct 540
|||||
Db 481 GTAGTCACGAGATGGTACCGAGCTCCTGAACTCTTGTGGGAGATACTCAGTATGGTTCT 540

QY 541 tcagtcgatatatgggctattggtgtgtgttttgcagagctcctgcagggccagccactg 600
|||||
Db 541 TCAGTCGATATATGGGCTATTGGTGTGTGTTTTGCAGAGCTCCTGACAGGCCAGCCACTG 600

QY 601 tggcctggaaaaatcagatgtggaccaactttatctgataatcagaaacactagtagagacg 660
|||||
Db 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACT----- 650

QY 661 gggtttcgcatgttgaccagggtgtgtctcgaaactcttgacgtcaaagtgcacctgcc 720
|||||
Db 651 ----- 650

QY 721 gtagcctctcaaaagtgtctggaattacaggaataataatcccaagacatcaatcaacttt 780
|||||
Db 651 -----AGGAAAATTAATCCCAAGACATCAATCAATCTTT 684

QY 781 aaaaagtaacgggttttccatggcatcagtcacgtacacctgcagcagaagacatggaaaacttt 840
|||||
Db 685 AAAAGTAAACGGGTTTTCCATGGCATCAGTATACCTGAGCCAGAGACATGGAAACTCTT 744

QY 841 gaggaanaagttctcagatgttcatcctgtggtctggaacttcattcatgaagggtgtctgaag 900
|||||
Db 745 GAGGAAAAGTCTCAGATGTTTCATCCTGTGGCTCTGAACTTTCATGAAGGGGTGTCTGAAG 804

QY 901 atgaatccagatgacagattaacctgttcccactcctggagagctcctactttgatct 960
|||||
Db 805 ATGAATCCAGATGACAGATTAACTGTTCCTCAACTCCTGGAGAGCTCCTACTTTGATTCT 864

QY 961 tttcaagaggcccaaatataaagaaaaagcacgtaatgaaggaagaaacagagacgcca 1020
|||||
Db 865 TTTCAGAGAGGCCCAAAATTAAGAGAAAAGCACGTAATGAAGGAAGAAACAGAGACGCCAA 924

QY 1021 caggtacttcgctcaaaaagt 1041
|||||
Db 925 CAGGTACTTCCGCTCAAAAGT 945
```

```
RESULT 4
AX107716 972 bp DNA PAT 30-APR-2001
LOCUS
DEFINITION Sequence 5 from Patent WO0123579.
ACCESSION AX107716
VERSION AX107716.1 GI:13923199
```


Db	301	GGAGTGATCAAAAGCGTATATATGGCAACACATTCAAGCTCTTAATTTCTGTCATATACAT	360
QY	361	aactgtattcacagagatataaaacctgaaaaatcttaataactaagcaaggaataatc	420
Db	361	AACGTATTACAGAGATATAAAACCTGAAAATATTTCTAATAACTAAGCAAGGAATAATC	420
QY	421	aagatttgacttcgggttgccacaaattctga-----	454
Db	421	AAGATTGTGACTTCGGGTTTGCACAAATTTCTGAGTAATTCACATGTTGGGAGAAATTGAC	480
QY	455	-----ttccaggagatgcctacacogattatgtagct	486
Db	481	TTACCTGATCTGATTGATGCATTTGCAAAATTTCTGAGTAATTCACATGTTGGGAGAAATTGAC	540
QY	487	acgagatggtacagagctcctgaaacttcttgtgggagatactcagtatggtcttcagtc	546
Db	541	ACGAGATGGTACCGAGCTCCTGAACTTCTTGTGGGAGATACTCAGTATGGTTCTTTCAGTC	600
QY	547	gatatatggctattggttggttttttgcagagctcctgcagagccagccactgtggcct	606
Db	601	GATATATGGGCTATTGGTTGTGTTTTTGCAAGACTCCTGCACAGCCAGCCACTGTGGCCT	660
QY	607	ggaaaatcagatgtggaccaactttatctgataatcagaacactagtagagacggggttt	666
Db	661	GGAAAATCAGATGTGGACCAACTTTATCTGATAATTCAGAACACT-----	704
QY	667	cgcatgttgaccagggtggtctcgaaactttgacgtcaagtgatccacctgccgtagcc	726
Db	705	-----	704
QY	727	tctcaaagtgctggaattacagggaataataatcccaagacatcaatcaatctttaaaagt	786
Db	705	-----AGGAAAATTAACTCCCAAGACATCAATCAATCTTTAAAAAGT	744
QY	787	aacggggttttccatgcagtcacgtacctgagccagaagacatggaaaactcttgaggaa	846
Db	745	AACGGGTTTTCCATGGCATCAGTATACCTTGAGCCAGAGACATGGAAAACCTCTTGAGGAA	804
QY	847	aagttctcagatgttccatcctgtggctctgaacttcagaggggtgctggaagatgaat	906
Db	805	AAGTTCTCAGATGTTTCATCCTGTGGCTCTGAACTTCATGAAGGGGTGCTGTAAGATGAAT	864
QY	907	ccagatgacagattaacctgtttcccaactcctggagagctcactcttgattcttttcaa	966
Db	865	CCAGATGACAGATTAACTGTTCCCAACTCCTGGAGAGCTCCTACTTTGATTCTTTTCAA	924
QY	967	gagggccaaattaaaaaagcacgtaataaggaagaaacagaaagacgacgccaacag	1023
Db	925	GAGGCCCAAATTAAAGAAAAGCACGTAATGAAGGAAGAAACAGAGACGCCAACAG	981
RESULT 6			
AX107724	AX107724	1819 bp	DNA
LOCUS	Sequence 13 from Patent WO0123579.	PAT	30-APR-2001
DEFINITION	AX107724		
ACCESSION	AX107724.1	GI:13923203	
VERSION			
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 1819)		
AUTHORS	Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.		
TITLE	Human kinase proteins and polynucleotides encoding the same		
JOURNAL	Patent: WO 0123579-A 13 05-APR-2001;		
FEATURES	Lexicon Genetics Incorporated (US)		
source	Location/Qualifiers		
	1. .1819		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		

BASE COUNT		539	a	368	c	406	g	505	t	1 others	
ORIGIN											
Query Match		68.1%; Score 709; DB 6; Length 1819;									
Best Local Similarity		77.0%; Pred. No. 1.1e-156;									
Matches 1025; Conservative		0; Mismatches 0; Indels 306; Gaps 1;									
QY	1	atgaaaaagtatgaaaaatagctaaagctggagaagggtcttattatgggggttgattccaaa	60								
Db	138	ATGGAAAAGTATGAAAAATTAGCTAAGACTGGAGAGGGTCTTTATGGGGTTGTATTCAAA	197								
QY	61	tcagaaaaacaaacacctctggacaagtagctgtttaaaaaatttgtggaatctgaagat	120								
Db	198	TGCAGAAACAAACCTCTGGACAAGTAGTACGTGTTAAAAAATTTGTGGAATCTGAAGAT	257								
QY	121	gatcctgttgttaagaaaaatagcactaagagaaaaatcgtatgttgaagcaattaaaaacat	180								
Db	258	GATCCTGTTGTTAAGAAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAAAAACAT	317								
QY	181	ccaaatcttgtgaacctcatcaggggtttcaggagagaaaaaggaaaaatgcatttagttttt	240								
Db	318	CCAAATCTTGTGAACCTCATCGAGGTTCAGGAGAAAAAGGAAAAATGCATTTAGTTTTT	377								
QY	241	gaatactgtgatcatcacacttttaaatgagctggaaagaaaccccaaatggagtgtgctgat	300								
Db	378	GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCCAATGGAGTTGCTGAT	437								
QY	301	ggagtgatcaaaaagcgattattatggcaaacacttcaagctcttaatttctgtcatatacat	360								
Db	438	GGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTCTATATACAT	497								
QY	361	aactgtattcacagagatataaaacctgaaactgaaataattcttaataacgaaggaataatc	420								
Db	498	AACTGTAATTCACAGAGATATAAAACCTGAAAAATATTCTAATAACTAAGCAAGGAATAATC	557								
QY	421	aagatttggacttcgggtttgcacaaaattctg-----	453								
Db	558	AAGATTTGTGACTTCGGGTTTGCACAAATTTCTGAGTTGGACTTCATCTTTCTCTGGTGCC	617								
QY	454	-----	453								
Db	618	TCCTTGATTGGCTTAATAGTTGACCTTCTGAAATTTCTTTTCTGCCAAATTCAGAGATTTT	677								
QY	454	-----	453								
Db	678	CTCCTGGCTTGGATCCATTGCTGACACAGTGTTCACCATGGGGCCCCAGGCTCATCTCGA	737								
QY	454	-----	453								
Db	738	ACTTCTGGCCTCAAGTGATCTTCCACCTCGGCCCTCCCAAAGTGTCTGGATTGCCAAGTGTG	797								
QY	454	-----	453								
Db	798	AGCCACCGTGCCAGCCAGATTTTTCAAACAATAACTACTGAGAGCTCACAGATTGTTT	857								
QY	454	-----attccaggagatgcctacacc	474								
Db	858	TTAGTGGGAACACAAATTTTCGAACAAATTTCTTGAGAACGCATTCAGGAGATGCCACACC	917								
QY	475	gattatgtagctacgagatgggtaccgagctcctgaacttcttgtgggagatactcagtat	534								
Db	918	GATTATGTAGTACGAGATGGTACCGAGCTCCTGAACCTTCTTGTGGAGATACTCAGTAT	977								
QY	535	ggttctcagtcgatataatgggctattgtgtgtttttgcagagctcctgcagggccag	594								
Db	978	GGTTCCTCAGTCGATATATGGGCTATTGGTGTGTTTTCGAGAGCTCCTGACAGGCCAG	1037								
QY	595	ccactgtggcctggaaaaatcagatgtggaccaactttatctgataatcagaacactagta	654								
Db	1038	CCACTGTGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGTA	1097								
QY	655	gagacggggttcccatgttgtgaccagggtggtctcgaactcttgacgtcaagtgatcca	714								

Db	1098		GAGACGGGGTTTCGCCATGTTGACCAAGGCTGGTCTCGAACTCTTGACGTCRAAGTGATCCA	1157
QY	715		cctgcgtagcctctcaaaagtctggaattacaggaaaaattaatcccaagacatcaatca	774
Db	1158		CCTGCCGTAGCCTCTCAAAGTCTGGAATTACAGGAAAATTATCCCAAGACATCAATCA	1217
QY	775		atcttaaaaagtaacgggtttttccatggccatcatataccctgagccaggaagacatgaa	834
Db	1218		ATCTTTAAAGTAACGGGTTTTCATGGCATCAGTATACCTTGAGCCAGAGACATGGAA	1277
QY	835		actcttgaggaaaaagtctcagatgttcatcctgtggctctgaaactlcatgaagggtgt	894
Db	1278		ACTCTTGAGGAAAAGTCTCAGATGTTTCATCCTGTGGCTCTGAACCTTCAITGAAGGGGTGT	1337
QY	895		ctgaagatgaatccagatgacagattaacctgttcccaactcctggagagctcctacttt	954
Db	1338		CTGAAGATGAATCCAGATGACAGATTAACCTGTCCCAACTCCTGGAGAGCTCCTACTTT	1397
QY	955		gattctttcaagagggcccaattaaaaagaaagcacgtaatgaagggaagaaacagaaga	1014
Db	1398		GATTCTTTCAAGAGGCCCAAAATTAAGGAAAAGCACGTAATGAAGGAAGAAACAGAAGA	1457
QY	1015		cgccaacacagt 1025	
Db	1458		CGCCAACAGGT 1468	
RESULT	7			
AX056404				
LOCUS	AX056404	911 bp	DNA	
DEFINITION	Sequence 48 from Patent WO0073469.		PAT	13-JAN-2001
ACCESSION	AX056404			
VERSION	AX056404.1	GI:12229111		
KEYWORDS	.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
TITLE	Plowman,G.D., Martinez,R., Whyte,D. and Sudersanam,S.			
JOURNAL	Patent: WO 0073469-A 48 07-DEC-2000;			
FEATURES	Sugen, Inc. (US)			
source	Location/Qualifiers			
	1..911			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
BASE COUNT	293 a	182 c	178 g	258 t
ORIGIN				
Query Match	51.6%;	Score 537;	DB 6;	Length 911;
Best Local Similarity	87.0%;	Pred. No. 2.4e-116;		
Matches	643;	Conservative 0;	Mismatches 0;	Indels 96;
				Gaps 1;
QY	285	aaatggagttgctgatggagtgatcaaaagcgtattatggcacaacacttcaagctcttaa	344	
Db	1	AAATGGAGTTGCTGATGGAGTGCATCAAAAGCGTATTATGGCAAACACTTCAAGCTCTTAA	60	
QY	345	ttctgtcatatacataaactgtattcacagagatataaaacctgaaaaattcttaataac	404	
Db	61	TTTCTGTATATACATAAAGTGTATTTCACAGAGATATAAAACCTGAAAATATTCTAATAAC	120	
QY	405	taagcaagggaataatcaagatttgtgacttcgggtttgcaaaaattctgattccagaga	464	
Db	121	TAAGCAAGGAATAATCAAGATTGTGACTTCGGGTTTGCACAAAATTCGATTCACAGAGA	180	
QY	465	tgcctacacgattatgtagctacagatggtaccgagctcctgaaacttcttgtggaga	524	
Db	181	TGCTACACCGATTATGTAGCTACGAGATGGTACCGAGCTCCTGAACCTCTTGTGGGAGA	240	
QY	525	tactcagtatggttcttcagtcgatataatgggctattgtgtgtttttgcagagctcct	584	

Db	241		TACTCAGTATGGTTCTTCAGTCGATATATGGGCTATGGTTGTGTTTTTGCAGAGCTCCT	300
QY	585		gacagggcagccactgtggcctggaaaatcagatgtggaccacttttatctgataatcag	644
Db	301		GACAGGCCAGCCACTGTGGCTGGAAATCAGATGTGGACCAACTTTATCTGATAATCAG	360
QY	645		aacactagtagagacggggtttcgccatggtgaccaggctggtctcgaactcttgacgtc	704
Db	361		AACACT-----	366
QY	705		aagtgatccacctgccgtagcctctcaaaagtgtggaattacaggaaaaataatcccaag	764
Db	367		-----AGGAAATTAATCCCAAG	384
QY	765		acatcaatcaatctttaaaagtaacgggtttttccatggcatcagtatacctgagccaga	824
Db	385		ACATCAATCAATCTTTAAAAAGTAACGGGTTTTCATGGCATCAGTATACCTGAGCCAGA	444
QY	825		agacatggaactcttgaggaaaaagtctcagatgttcatcctgtggtctctgaacttcat	884
Db	445		AGACATGGAACCTCTTGAGGAAAAAGTCTCAGATGTTTCATCCTGTGGCTCTGAACCTCAT	504
QY	885		gaagggtgtctgaagatgaatccagatgacagattaacctgttcccaactcctggagag	944
Db	505		GAAGGGTGTCTGAAGATGAATCCAGATGACAGATTAACTGTTCCCAACTCCTGGAGAG	564
QY	945		ctcctactttgattctttcaagagggcccaattaaaaagaaagcacgtaataagaggaag	1004
Db	565		CTCCTACTTGTATTCTTTCAAGAGGCCCAAAATTAAGGAAAAAGCACGTAATGAAGGAAG	624
QY	1005		aaacagaagacgccaacag 1023	
Db	625		AAACAGAAGACGCCAACAG 643	
RESULT	8			
AX107712				
LOCUS	AX107712	561 bp	DNA	
DEFINITION	Sequence 1 from Patent WO0123579.		PAT	30-APR-2001
ACCESSION	AX107712			
VERSION	AX107712.1	GI:13923197		
KEYWORDS	.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
TITLE	Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and			
JOURNAL	Sands,A.T.			
FEATURES	Human kinase proteins and polynucleotides encoding the same			
source	Patent: WO 0123579-A 1 05-APR-2001;			
	Lexicon Genetics Incorporated (US)			
	Location/Qualifiers			
	1..561			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
BASE COUNT	189 a	87 c	112 g	173 t
ORIGIN				
Query Match	43.7%;	Score 454.4;	DB 6;	Length 561;
Best Local Similarity	99.8%;	Pred. No. 5.7e-97;		
Matches	455;	Conservative 0;	Mismatches 1;	Indels 0;
				Gaps 0;
QY	1	atggaaaagtatgaaaaattagctaagactggagaagggtcttatggggttgattcaaa	60	
Db	1	ATGGAAAAGTATGAAAATTAGCTAAGACTGGAGAAGGGTCTTATGGSGTTGTATTCAA	60	
QY	61	tcagaaaacaaaacctctggacaagtagtacctgttaaaaaatttgggaactctgaagat	120	
Db	61	TGCAGAAACAAAACCTCTGGACAAGTAGTAGCTGTAAATAATTTGTGGAATCTGAAGAT	120	

QY	121	gatcctgtttaaagaaaatagcactaagagaaaatacgtatgttgaagcaattaaaaacat	180
Db	121	GATCCTGTTGTTAAGAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAAAAACAT	180
QY	181	ccaaatcttgtgaacctcatcgaggtgttcaggagaaaaaggaaaaatgcatttagtttt	240
Db	181	CCAAATCTTGTGAACCTCATCGAGGTGTTCAAGGAGAAAAAGGAAAATGCATTAGTTTTT	240
QY	241	gaatacttgtatcacatcacttttaaatgagctgggaagaaacccaaaatggagttgctgat	300
Db	241	GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAAATGGAGTTGCTGAT	300
QY	301	ggagtgatcaaaagcgtattatggcacaacacttcaagctcttaattcttgcataatacat	360
Db	301	GGAGTGATCAAAAGCGTATTATGGCAACACACTTCAAGCTCTTAATTCTGTGCATATACAT	360
QY	361	aactgtattcacagagatataaaacctgaaaatattcttaataaactaagcaaggaataatc	420
Db	361	AACTGTATTCACAGAGATATAAAACCTGAAAATATTCTAATAACTAAGCAAGGAATAATC	420
QY	421	aagatttgtgacttcgggtttgcacaaaattctgatt	456
Db	421	AAGATTTGTGACTTCGGGTTTGACAAAATTCTGAGT	456
RESULT	9		
LOCUS	AX107718	594 bp	DNA
DEFINITION	Sequence 7 from Patent WO0123579.	PAT	30-APR-2001
ACCESSION	AX107718		
VERSION	AX107718.1	GI:13923200	
KEYWORDS	human.		
SOURCE			
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 594)		
AUTHORS	Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.		
TITLE	Human kinase proteins and polynucleotides encoding the same		
JOURNAL	Patent: WO 0123579-A 7 05-APR-2001; Lexicon Genetics Incorporated (US)		
FEATURES	Location/Qualifiers		
source	1..594		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	203 a	94 c	115 g
ORIGIN	182 t		
Query Match	43.7%;	Score 454.4;	DB 6; Length 594;
Best Local Similarity	99.8%;	Pred. No. 5.7e-97;	
Matches	455;	Conservative 0;	Mismatches 1; Indels 0; Gaps 0;
QY	1	atggaaaagtatatgaaaaattagctaagactggagaagggtcttatggggttgattcaaa	60
Db	1	ATGGAAGAATATGAAAAATTAGCTAAGACTGGAGAAGGGTCTTATGGGGTTGTATCAAA	60
QY	61	tgcagaaaacaaacacctctggacaagttagtgtgttaaaaaatttgtgaaatctgaagat	120
Db	61	TGCAGAAACAAACACCTCTGGACAAGTAGTAGCTGTIAAAAAATTTGTGGAATCTGAAGAT	120
QY	121	gatactgtttaaagaaaatagcactaagagaaaatacgtatgttgaagcaattaaaaacat	180
Db	121	GATCCTGTTGTTAAGAAAATAGCACTAAGAGAAAATACGTATGTTGAAGCAATTAAAAACAT	180
QY	181	ccaaatcttgtgaacctcatcgaggtgttcaggagaaaaaggaaaaatgcatttagtttt	240
Db	181	CCAAATCTTGTGAACCTCATCGAGGTGTTCAAGGAGAAAAAGGAAAATGCATTAGTTTTT	240
QY	241	gaatacttgtatcacatcacttttaaatgagctgggaagaaacccaaaatggagttgctgat	300
Db	241	GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAAATGGAGTTGCTGAT	300
QY	421	aagatttgtgacttcgggtttgcacaaaattctgatt	456
Db	421	AAGATTTGTGACTTCGGGTTTGACAAAATTCTGAGT	456

QY	301	ggagtgatcaaaagcgtattatggcaaacacacttcaagctcttaattcttgcataatacat	360
Db	301	GGAGTGATCAAAAGCGTATTATGGCAACACACTTCAAGCTCTTAATTCTGTGCATATACAT	360
QY	361	aactgtattcacagagatatataaacctgaaaatattcttaataaactaagcaaggaataatc	420
Db	361	AACTGTATTACAGAGATATAAAACCTGAAAATATTCTAATAACTAAGCAAGGAATAATC	420
QY	421	aagatttgtgacttcgggtttgcacaaaattctgatt	456
Db	421	AAGATTTGTGACTTCGGGTTTGACAAAATTCTGAGT	456
RESULT	10		
LOCUS	AX056405	2615 bp	DNA
DEFINITION	Sequence 49 from Patent WO0073469.	PAT	13-JAN-2001
ACCESSION	AX056405		
VERSION	AX056405.1	GI:12229112	
KEYWORDS	Murinae gen. sp.		
SOURCE	Murinae gen. sp.		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae.		
REFERENCE	1 (bases 1 to 2615)		
AUTHORS	Plowman,G.D., Martinez,R., Whyte,D. and Sudersanam,S.		
TITLE	Protein kinases		
JOURNAL	Patent: WO 0073469-A 49 07-DEC-2000; Sugen, Inc. (US)		
FEATURES	Location/Qualifiers		
source	1..2615		
	/organism="Murinae gen. sp."		
	/db_xref="taxon:39108"		
BASE COUNT	742 a	580 c	585 g
ORIGIN	708 t		
Query Match	43.5%;	Score 453.2;	DB 6; Length 2615;
Best Local Similarity	73.1%;	Pred. No. 1.2e-96;	
Matches	658;	Conservative 0;	Mismatches 143; Indels 99; Gaps 2;
QY	135	gaaaaatagcactaagagaaaatcgtatgttgaagcaattaaacaatccaaatcttgtgaa	194
Db	1	GAAAATAGCCCTGCGGGAATCCGTATCTGCTGAAG--TTGAAACACCCCAACCTCGTGAA	57
QY	195	cctcatcgaggtgttcaggagaaaaaggaaaatcatttagttttgaatactgtgatca	254
Db	58	CCTCATCGAGGTGTTCAGAAGAAAGAGAAAAGATGCATCTAGTTTGTAGTACTGTGATCA	117
QY	255	tacacttttaaatgagctgggaaagaaaccccaatggagttgctgatggagtgatcaaaag	314
Db	118	CACACTGTTAAACGAGCTGGAGAGAAACCCCAACGGAGTTTCTGTGATGGAGTGATTAAAG	177
QY	315	cgtattatggcaaacacttcaagctcttaatttctgtcatatacacataaactgtattcacag	374
Db	178	TGTGCTATGGCAAAACCTTCAAGCCCTTAACTTCTGTCTACAAAGCACAATTTGTATTATCG	237
QY	375	agatataaaacctgaaaaatattcttaataaactaagcaaggaataatcaagatttgtgactt	434
Db	238	GGATGTAAACCTGAAAACATCCTTAATAACCAAGCAAGGGGATGATAAAGATTGTGACTT	297
QY	435	cggtttgcaaaaatcttgattccaggagatgcctacacccgattatgtagctacgagatg	494
Db	298	TGGATTGACGAATTCTTAATCCAGGAGACGCCCTACACAGACTATGTTGCCACCAGGTG	357
QY	495	gtaccgagctcctgaacttcttgtgggagatactcagtatggttcttcagtcgatatatg	554
Db	358	GTACCGAGCCCCGAACTTCTCGTGGGAGACACGGAAGTACGGTTCCCTCTGTAGACGTGTG	417
QY	555	ggctattggttgtgtttttgcagagctcctgcagaggccagccactgtggcctggaaaaatc	614
Db	418	GGCCGTCGGCTGTGTTTTTGCAGAGCTCTCTGACGGGTGAGCCCACTCTGGCCGGGAAAAATC	477

Qy 615 agatgtggaccaactttatctgataaatcagaacacactagtagagacaggggtttcgccatgt 674
Db 478 CGACGTGGACCAGCTTTACCTGATCATCAGGACGTTG-----514
Qy 675 tgaccaggctggtctcgaactcttgacgtcaagtgatccacactgccgtagcctctctcaag 734
Db 515 -----514
Qy 735 tgctggaattacaggaaaaattaatcccaagacatacaatcaatcttltaaaagtaacgggtt 794
Db 515 -----GGGAAGCTGATTCCAAGACACACAGTCTATCTTTAGGAGTAACCGATT 561
Qy 795 ttccatggcatcagtataacctgagccagaagacatggaaaactcttgagaaaaagttctc 854
Db 562 TTTCCGCGGCATCAGCATACCTGAACGAGGACATGGAGACTCTTTGAAGAAAAATTC 621
Qy 855 agatgttcctcctgtggtcctgaaacttcataaggggtgtctgaaagatgaatccagatga 914
Db 622 AATGTTTCAGCTGTGGCTTTAAGTTTCATGAAGGATGCCTGAAGATGAATCCTGATGA 681
Qy 915 cagattaaacctgttcccaactcctggagagctcctactcttctgattcttctcaagagccca 974
Db 682 GAGCTGACCTGTGCCAGCTGTGCACAGTGCCTACTTTGAGTCTTTTCAAGAGGATCA 741
Qy 975 aattaaaaaagaaagcacgtaatgaaggaagaacagaaagacgccaacaggctacttccgct 1034
Db 742 AATGAAAAGAAAAGCCCGCAGTGAGGGGAGAGCGGAAGCGCCAGCAGCAATCAACTGCT 801

RESULT 11
HSSTHPKB HSSTHPKB 1363 bp mRNA PRI 06-FEB-1997
LOCUS H.sapiens mRNA KKIALRE for serine/threonine protein kinase.
DEFINITION X66358
ACCESSION X66358
VERSION X66358.1 GI:36614
KEYWORDS serine/threonine protein kinase.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1363)
AUTHORS Meyerson,M.L.
TITLE Direct Submission
JOURNAL Submitted (12-MAY-1992) M.L. Meyerson, Massachusetts General
Hospital, Cancer Center, Bldg 149, 13th Street, Charleston MA
02129, USA
REFERENCE 2 (bases 1 to 1363)
AUTHORS Meyerson,M., Enders,G.H., Wu,C.L., Su,L.K., Gorka,C., Nelson,C.,
Harlow,E. and Tsai,L.H.
TITLE A family of human cdc2-related protein kinases
JOURNAL EMBO J. 11 (8), 2909-2917 (1992)
MEDLINE 92347325
FEATURES Location/Qualifiers
source 1..1363
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="HeLa"
/cell_line="HeLa"
/clone_lib="HeLa"
CDS join(214..669,X66359.1:1..54,670..1290)
/note="KKIALRE coding sequence with 54 bp insertion"
/codon_start=1
/pseudo
/label=sthpkin_CDS
/product="serine/threonine protein kinase"
/db_xref="PID:e49306"
214..1290
/codon_start=1
/product="serine/threonine protein kinase"
/protein_id="CAA47002.1"
CDS /db_xref="GI:36615"
/db_xref="SWISS-PROT:Q00532"

/translation="MMEKYEKIGKIGESYGVVEKCRNRDTGQIVAIAKKFLESEDDPV
IKKIALREIRMLKQLAKHPNLVNLLEVFERRKRRLHLVFEYCDHTVLHEDRYQRGVPEH
LVKSITWOTLQAVNFCCHKNCIHRDVKPENILITKHSVILKCDFGFARLLTGPSDYT
DYVATRWYRSPPELLLVGDTQYGPVVDVWAIICVFAELLSGVPLWPKSDVDQLYLIRKT
LGDLIQRHQVFSTNQYFSGVKIPDPEDMEPLELKFNPISYPALGLLKGCLHMDPTFR
LTCEQLLHHPYFENIRIEDLAKEDHKPTRKTLRKSRKHKCFETETSKLQYLPLQLTGSS
ILPALDNKKYYCDTKKINIRFPNI"
BASE COUNT 412 a 308 c 309 g 334 t
ORIGIN
Query Match 33.0%; Score 343.8; DB 9; Length 1363;
Best Local Similarity 71.2%; Pred. No. 5.4e-71;
Matches 469; Conservative 0; Mismatches 187; Indels 3; Gaps 1;
Qy 1 atggaagaagtatgaaaaattagctaagactggagaaggtcttatggggttgtattcaaa 60
Dy 217 ATGGAGAAGTATGAAAAAATTGGGAAAATTGGAGAAGGATCCTATGGAGTTGTTTCAA 276
Qy 61 tgcagaacaacaaacacctctgtgacaaagttagctgtgttaaaaaatttgtggaatctgaagat 120
Dy 277 TGTAGAAACAGGGACACAGGGTTCAGATGTGGCCATCAAGAAGTTTCTTGAATCAGAAGAT 336
Qy 121 gatcctgtgttaagaataatagcactaaagagaaatacgtatgttgaaagcaattaaaaacat 180
Dy 337 GACCCTGTGATTAAGAAAAATTGCCCTTCGGGAAATCCGAATGCTCAAGCAACTCAAGCAT 396
Qy 181 ccaaatcttgtgaacctcatcgaggtgttcagagagaaaaaggaataatgcatttagtttt 240
Dy 397 CCCAACCTTGTTAACCTCCTGGAGTCTTCAGAGGAAACGGAGGCTTCACTGCTGTGTTT 456
Qy 241 gaatactgtgatcatatacaacttttaaatgagctggaaagaaaccccaatggagttgctgat 300
Dy 457 GAATATTGTGACCACACAGTTCCTCATGAGTTGGACAGATACCAAGAGGGGTACAGAA 516
Qy 301 ggagtgatcaaaaagcgtattatggcaaacacttcaagcttcttaattttgtcatatacat 360
Dy 517 CATCTCGTGAAGAGCATAAATTGGCAGACACTGCAAGCTGTAAATTTTGGCCATAAACAC 576
Qy 361 aactgtattcacagagatataaaacacctgaaaaattcttaataactaaagcaaggaataatc 420
Dy 577 AATTGCATACATAGACACGTGAAGCCAGAAAAATATCCTCATACGAAACATTCCTGTGATT 636
Qy 421 aagatttgtacttcgggtttgcacaaaattctgatt---ccaggagatgcctacaccgat 477
Dy 637 AAGCTTTGTGACTTTGGATTTCCTCGGCTTTTGACTGGACCGAGTACTACTATACAGAC 696
Qy 478 tatgtagctacgagatggtaccgagctcctgaacttctgtgggagatactcagtagtgt 537
Dy 697 TACGTGCTTACCAGTGGTACCGCTCCCGCTCCCTGAGCTGCTGGTGGGGACACGCGTACGGC 756
Qy 538 tcttcagtcgatataatgggctattgggtgtgttttttcagagagctcctgacaggccagcca 597
Dy 757 CCCCCGGTGGATGTTGGGCAATTGGCTGTGCTCTTTGCTGAGCTGCTGTCAGGAGTGCCT 816
Qy 598 ctgtggcctggaaaaatcagatgtggaccacacttttatctgataatacagaacactagtaga 656
Dy 817 CTGTGCCCAGGAAAAATCGGATGTGGATCAGCTGTAICTGATTAGGAAGACCTTGGGGGA 875

RESULT 12
AB029045
LOCUS Oryctolagus cuniculus mRNA for Ser/Thr kinase KKIAMRE, complete
DEFINITION cds.
ACCESSION AB029045
VERSION AB029045.1 GI:6624278
KEYWORDS Ser/Thr kinase KKIAMRE.
SOURCE Oryctolagus cuniculus male brain cDNA to mRNA.
ORGANISM Oryctolagus cuniculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
REFERENCE 1 (sites)

AUTHORS Gomi,H., Sun,W., Finch,C.E., Itohara,S., Yoshimi,K. and Thompson,R.F.
TITLE Learning induces a cdc2-related protein kinase, kkiamre
JOURNAL Unpublished (1999)
REFERENCE 2 (bases 1 to 3153)
AUTHORS Gomi,H. and Itohara,S.
TITLE Direct Submission
JOURNAL Submitted (16-JUN-1999) to the DDBJ/EMBL/GenBank databases. Hiroshi Gomi, The Institute of Physical and Chemical Research (RIKEN), Brain Science Institute (BSI); 2-1 Hiroosawa, Wako, Saitama 351-0198, Japan (E-mail:hgomi@brain.riken.go.jp, Tel:+81-48-467-9724, Fax:+81-48-467-9725)

FEATURES
source
1. .3153
/organism="Oryctolagus cuniculus"
/db_xref="taxon:9986"
/sex="male"
/tissue_type="brain"
322. .2022
/note="corresponds to mouse KKIAMRE-beta"
/codon_start=1
/product="Ser/Thr kinase KKIAMRE"
/protein_id="BAA88508.1"
/db_xref="GI:6624279"
/translation="MEKYENLGLVGEISYGMVKCRNKDSGRIVAIAKKFLESDDDKMV KKIAMREIKLLQLRHENLVNLEVCCKKKRWYLVFEFVDHTILDDLELFPNGLDDQV VOKYLFQILINGFCHSHNIHRDIKPENILVSQGVVKLCDFGFARTLAAPGEVYTD YVATRWYRAPELLVGDVYKGRAVDVWAIGCLVTEMLNGEPLFPGSDSIDQLYLMRCL GNLI PRHOELFYKNPVFAGVRLPEIKSEPLRYKLVSEVVIDLAKKCLHVDPDKRP FCAELLHHDFQMDGFAERFSQELQMKVOKDARNISLSKSKSNRKEKEKEDSLGEER KTLVVDQTNVDSKFKIKGSKIDGKVDKNRAAVSMTVPSPSHIKAVPSTSLR DCSNVSDHTRNPQMAIPPLTHNLSAVAGINSMGMTIPGVQSYRVDEKTKYKCIPIV KPNKHSPSGIYNNMVTTSVSEKNLLQANKKRGEYSKTDVRLPELNYNHLPELRALEG IARNSRLIRKENKILSESRIPSLAAIDLHTPNIAVHVQVSGSPSLSDGSEADSPWMEHQH "

BASE COUNT 939 a 577 c 777 g 860 t
ORIGIN
Query Match 24.2%; Score 251.8; DB 4; Length 3153;
Best Local Similarity 62.9%; Pred. No. 2.2e-49;
Matches 407; Conservative 0; Mismatches 237; Indels 3; Gaps 1;

QY 1 atgaaaaagtatgaaaaattagctaagactggagaagggtcttatggggtgtattcaaa 60
||||| ||||| || | | ||||| ||||| ||||| || | | |
Db 322 ATGGAATAATGAGAACCTTGGATTGGTTGGAGAGGGAGCTATGGAATGGTGAAG 381

QY 61 tgcagaaacaaacacctctggacaagtagtagctgttaaaaaatttgtggaatctgaagat 120
|| | || |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | ||||
Db 382 TGTAGGAATAAGATAGTGAAGAAATTGTGGCCATCAAGAACTCTCTAGAAAGTGATGAT 441

QY 121 gatcctgttgttaagaaaaatagcactaagagaaaaatcgttatgtgaagcaattaaaaacat 180
|| | ||||| ||||| || | ||||| | | ||||| | | ||||| | | ||||| | | ||||
Db 442 GACAAATGGTTAAAAAATTGCTATGCGAGAAATCAAGTTACTAAAGCAACTGAGGCAT 501

QY 181 ccaaatcttgtgaacctcatcgaggtgttcaggagaaaaaggaaaaatgcatttagttttt 240
||||| ||||| || | ||||| || | ||||| | | ||||| | | ||||| | | |||||
Db 502 GAAATTTGGTGAATCTGTTGGAGGTGTGTAAAAAATAAACGATGGTACCTAGTCTTT 561

QY 241 gaactgtgatcacacacttttaaatgactggaaagaaacccaaatggagttgctgat 300
|||| | || | || | || | || | || | || | || | || | || | || | || | || | || |
Db 562 GAATTTGTCACCACACGATCTTGTATGACTTGGAACTCTTCCAAATGGACTAGATGAC 621

QY 301 ggagtgatcaaaacgctattatggcaaacacttcaagctcttaatttctgtcatatacat 360
|| | || | || | || | || | || | || | || | || | || | || | || | || | || |
Db 622 CAAGTAGTTCAAAAGTATTTGTTTCAGATATTAATGGAATTGGATTTTGTACAGTCAAC 681

QY 361 aactgtattcacagagatataaaacctgaaaaatattctaataactaagcaaggaataatc 420
|| | || | ||||| |||| | ||||| | | ||||| | | ||||| | | ||||| | | ||||
Db 682 AATATCATACATAGAGATATAAAGCCAGAGAAATATATTGGTCTCCAGCTCGGCGTTGTC 741

QY 421 aagatttgtacttcgggtttgcacaaaattctg---attccaggagatgcctacaccgat 477
|| | ||||| || | ||||| || | ||||| || | ||||| || | ||||| || | ||||| || |

Db 742 AAGTTATGTGATTTTGGATTTGCACGGACACTGGCAGCTCCCGGAGAGGTTTACACTGAT 801

QY 478 tatgtactacgagatggtaccgagctcctcgaacttcttgtgggagatactcagtatggt 537
||||| || | ||||| ||||| ||||| ||||| || | || | || | ||||| ||||| |||||
Db 802 TATGTGCAACTCGATGGTACAGAGCTCCAGAACTACTGGTTGGTGATGTCAAGTATGGC 861

QY 538 tcttcagtcgatatatgggctattgtgtgtgttttttgcagagactcctgcacagggccagcca 597
| | | || | ||||| ||||| ||||| || | | | || | || | || | || | || | || |
Db 862 AAAGCTGTGGATGTGTGGCCATGGTTGTCTGGTAACTGAAATGCTCATGTTGGGAACCC 921

QY 598 ctgtggcctggaataatcagatgtggaccaacttttatctgataataatcag 644
|||| | ||||| | | || | || | || | ||||| || | ||||| || | ||||| || | |||||
Db 922 CTGTTTCTCTGGAGACTCTGATATTGATCAGCTTTTATCTTATTATGAG 968

RESULT 13
AB029067
LOCUS AB029067 1518 bp mRNA ROD 17-DEC-1999
DEFINITION Mus musculus Kkm mRNA for Ser/Thr kinase KKIAMRE-delta, partial cds.
ACCESSION AB029067
VERSION AB029067.1 GI:6594327
KEYWORDS Ser/Thr kinase KKIAMRE-delta.
SOURCE Mus musculus (strain:C57BL/6) adult brain tissue lib:Uni-ZAP cDNA to mRNA.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)
AUTHORS Sassa,T., Gomi,H., Sun,W., Ikeda,T., Thompson,R.F. and Itohara,S.
TITLE The Murine KKIAMRE gene: variants, dual promoters, expression and chromosomal localization
JOURNAL Unpublished (1999)
REFERENCE 2 (bases 1 to 1518)
AUTHORS Sassa,T., Gomi,H. and Itohara,S.
TITLE Direct Submission
JOURNAL Submitted (16-JUN-1999) to the DDBJ/EMBL/GenBank databases. Takayuki Sassa, The Institute of Physical and Chemical Research (RIKEN), Brain Science Institute (BSI); 2-1 Hiroosawa, Wako, Saitama 351-0198, Japan (E-mail:tasasa@brain.riken.go.jp, Tel:81-48-467-9724(ex.7716), Fax:81-48-467-9725)

FEATURES
source
1. .1518
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/dev_stage="adult"
/tissue_lib="Uni-ZAP"
/tissue_type="brain"
530. .1518
/gene="Kkm"
530. .>1518
/gene="Kkm"
/note="cdc2-related kinase"
/codon_start=1
/product="Ser/Thr kinase KKIAMRE-delta"
/protein_id="BAA88429.1"
/db_xref="GI:6594328"
/translation="MEKYENLGLVGEISYGMVKCRNKDSGRIVAIAKKFLESDDDKMV KKIAMREIKLLQLRHENLVNLEVCCKKKRWYLVFEFVDHTILDDLELFPNGLDYQV VOKYLFQILINGFCHSHNIHRDIKPENILVSQGVVKLCDFGFARTLAAPGEVYTD YVATRWYRAPELLVGDVYKGRAVDIWAIGCLVIELMGQPLFPGESIDQLHHIMTCL GNLI PRHOELFYKNPVFAGVRLPEVKDAEAEPLSRYPKLPEAVISLAKKCLHIDPDK RPFCADLLRHDFQMDGFAERFSQELQLKIEKDARNNSLPKKSQKKKKKKKKK"

BASE COUNT 421 a 303 c 382 g 412 t
ORIGIN
Query Match 23.4%; Score 243.8; DB 10; Length 1518;
Best Local Similarity 61.8%; Pred. No. 1.6e-47;
Matches 405; Conservative 0; Mismatches 247; Indels 3; Gaps 1;
QY 1 atggaaaaagtatgaaaaattagctaagactggagaagggtcttatggggtgtattcaaa 60

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 15:42:23 ; Search time 234.75 Seconds
(without alignments)
3801.813 Million cell updates/sec

Title: US-09-671-050-9
Perfect score: 1041
Sequence: 1 atggaaaagtatgaaaaatt.....aggtacttcgcgtcaaaagt 1041

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :	N_Geneseq_1101.*
1:	/SIDS1/gcgdata/geneseq/geneseqn/NA1980.DAT.*
2:	/SIDS1/gcgdata/geneseq/geneseqn/NA1981.DAT.*
3:	/SIDS1/gcgdata/geneseq/geneseqn/NA1982.DAT.*
4:	/SIDS1/gcgdata/geneseq/geneseqn/NA1983.DAT.*
5:	/SIDS1/gcgdata/geneseq/geneseqn/NA1984.DAT.*
6:	/SIDS1/gcgdata/geneseq/geneseqn/NA1985.DAT.*
7:	/SIDS1/gcgdata/geneseq/geneseqn/NA1986.DAT.*
8:	/SIDS1/gcgdata/geneseq/geneseqn/NA1987.DAT.*
9:	/SIDS1/gcgdata/geneseq/geneseqn/NA1988.DAT.*
10:	/SIDS1/gcgdata/geneseq/geneseqn/NA1989.DAT.*
11:	/SIDS1/gcgdata/geneseq/geneseqn/NA1990.DAT.*
12:	/SIDS1/gcgdata/geneseq/geneseqn/NA1991.DAT.*
13:	/SIDS1/gcgdata/geneseq/geneseqn/NA1992.DAT.*
14:	/SIDS1/gcgdata/geneseq/geneseqn/NA1993.DAT.*
15:	/SIDS1/gcgdata/geneseq/geneseqn/NA1994.DAT.*
16:	/SIDS1/gcgdata/geneseq/geneseqn/NA1995.DAT.*
17:	/SIDS1/gcgdata/geneseq/geneseqn/NA1996.DAT.*
18:	/SIDS1/gcgdata/geneseq/geneseqn/NA1997.DAT.*
19:	/SIDS1/gcgdata/geneseq/geneseqn/NA1998.DAT.*
20:	/SIDS1/gcgdata/geneseq/geneseqn/NA1999.DAT.*
21:	/SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT.*
22:	/SIDS1/gcgdata/geneseq/geneseqn/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1041	100.0	1041	22	Human kinase cDNA
2	1025	98.5	1068	22	Human kinase cDNA
3	839	80.6	945	22	Human kinase cDNA
4	823	79.1	972	22	Human kinase cDNA
5	757	72.7	1083	22	Polynucleotide seq
6	709	68.1	1819	22	Human kinase cDNA
7	537	51.6	911	22	Novel protein kina
8	454.4	43.7	561	22	Human kinase cDNA
9	454.4	43.7	594	22	Human kinase cDNA
10	453.2	43.5	2615	22	Novel protein kina
11	251.8	24.2	1698	21	Rabbit KKIAMRE kin

12	251.8	24.2	3080	21	AAA29745	Rabbit KKIAMRE kin
13	202.6	19.5	1513	21	AAZ51208	Human lost in leuk
14	202.6	19.5	1773	22	AAS06724	Polynucleotide seq
15	191	18.3	1667	21	AAZ51207	Rat lost in leukae
16	191	18.3	1738	21	AAZ51206	Rat lost in leukae
17	175.2	16.8	1866	22	AAF44672	Novel protein kina
18	163.8	15.7	903	20	AAX07476	Pneumocystis carin
19	162	15.6	1476	21	AAZ37835	Cyclin-dependent k
20	162	15.6	1635	19	AAV71073	CDK2-green fluore
21	162	15.6	1635	19	AAV71074	Green fluorescent
22	160	15.4	1322	21	AAF15889	Human prostate can
23	134.8	12.9	1825	22	AAI58432	Human polynucleoti
24	133.2	12.8	1089	15	AAQ63490	Cyclin dependent p
25	128	12.3	1002	18	AAI64448	CDK1 protein. Can
26	128	12.3	1002	21	AAZ08264	Candida gene encod
27	127	12.2	6015	21	AAC75931	Human ORFX ORF1486
28	125.6	12.1	1788	22	AAI60218	Human polynucleoti
29	125.4	12.0	1310	21	AAC35344	Arabidopsis thalia
30	125.4	12.0	1899	22	AAF44678	Novel protein kina
31	124.6	12.0	1293	21	AAC36413	Arabidopsis thalia
32	123.8	11.9	1307	21	AAC51531	Arabidopsis thalia
33	123.4	11.9	1326	21	AAC43758	Arabidopsis thalia
34	116	11.1	1070	18	AAT64450	Zea mays DNA fragm
35	116	11.1	1070	21	AAZ08266	MOC1 protein. Can
36	115.4	11.1	1353	21	AAC46587	Candida gene encod
37	110.8	10.6	996	22	AAH46905	zea mays DNA fragm
38	107.2	10.3	936	22	AAF58252	cDNA encoding huma
39	107.2	10.3	936	22	AAF58254	Oligonucleotide D1
40	107.2	10.3	936	22	AAF58257	Oligonucleotide D1
41	107.2	10.3	936	22	AAF58259	Oligonucleotide D2
42	107.2	10.3	936	22	AAF58262	Oligonucleotide D2
43	107.2	10.3	938	22	AAF58255	Oligonucleotide D1
44	106.2	10.2	936	22	AAF58252	Oligonucleotide D1
45	106.2	10.2	936	22	AAF58254	Oligonucleotide D1

ALIGNMENTS

RESULT	1
AAD03816	
ID	AAD03816 standard; cDNA; 1041 BP.
XX	
AC	AAD03816;
XX	
DT	19-JUN-2001 (first entry)
XX	
DE	Human kinase cDNA #5.
XX	
KW	Human; kinase; gene therapy; bioreactor; mental disorder;
KW	biological disorder; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	1..1041
FT	/*tag= a
FT	/product= "Human kinase #5"
FT	/note= "The coding region does not include stop codon"
FT	/partial
XX	
PN	WO200123579-A1.
XX	
PD	05-APR-2001.
XX	
PF	27-SEP-2000; 2000WO-US26621.
XX	
PR	28-SEP-1999; 99US-0156511.
XX	
PA	(LEXI-) LEXICON GENETICS INC.
XX	
PI	Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
XX	

DR WPI; 2001-266166/27.
DR P-PSDB; AAE00494.
XX
PT New isolated human kinase polynucleotide useful for generating
PT antibodies, as reagents in diagnostic assays and for screening for
PT compounds useful for treating mental, biological or medical diseases
XX
PS Claim 1; Page 31-32; 38pp; English.
XX
CC The present sequence is a cDNA encoding novel human protein
CC (NHP) known as human kinase. The human kinases share structural
CC similarity with animal kinases, more particularly serine or
CC threonine protein kinases. Human kinase cDNA is useful for the
CC detection of mutant human kinase for the diagnosis of disease,
CC and also as a therapeutic. It is useful for screening drugs
CC effective in the treatment of symptomatic or phenotypic
CC manifestations perturbing the normal function of NHP in the
CC body. The NHP nucleotide sequences are useful for generation of
CC antibodies, as reagents in diagnostic assays, for the
CC identification of other cellular gene products related to human
CC kinases, and as reagents in assays for screening compounds that
CC are useful for treating mental, biological or medical disorders.
CC NHP oligonucleotides are used as probes. The labelled NHP probes
CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
XX modulation of NHP expression.
SQ Sequence 1041 BP; 344 A; 190 C; 228 G; 279 T; 0 other;

Query Match 100.0%; Score 1041; DB 22; Length 1041;
Best Local Similarity 100.0%; Pred. No. 1.1e-268;
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atggaaaaagtatgaaaaattagctaagactggagagggtcttatggggtgtattcaaa 60
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
1 atggaaaaagtatgaaaaattagctaagactggagagggtcttatggggtgtattcaaa 60
QY 61 tgcagaaaaacaaacctctggacaagttagctgttaaaaaatttgggaatctgaaagat 120
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
61 tgcagaaaaacaaacctctggacaagttagctgttaaaaaatttgggaatctgaaagat 120
QY 121 gatcctgttgttaagaaaaatagcactaagagaaaatcgtatgttgaagcaattaaacat 180
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
121 gatcctgttgttaagaaaaatagcactaagagaaaatcgtatgttgaagcaattaaacat 180
QY 181 ccaaatcttgtgaacctcatcgagggtgttcagagaaaaaggaaaaatgcatttagtttt 240
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
181 ccaaatcttgtgaacctcatcgagggtgttcagagaaaaaggaaaaatgcatttagtttt 240
QY 241 gaatactgtgatcatcacacttttaaatgagctggagaaagaaacccaaatggagttgctgat 300
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
241 gaatactgtgatcatcacacttttaaatgagctggagaaagaaacccaaatggagttgctgat 300
QY 301 ggagtgatcaaaaagcgtattatggcaaacacttcaagctcttaatttctgtcatatacat 360
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
301 ggagtgatcaaaaagcgtattatggcaaacacttcaagctcttaatttctgtcatatacat 360
QY 361 aactgtattcacagagatatataaacctgaaaatatcttaataactaagcaaggaataatc 420
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
361 aactgtattcacagagatatataaacctgaaaatatcttaataactaagcaaggaataatc 420
QY 421 aagatttgtgacttcgggtttgcacaaaattctgattccaggagatgcctacaccgattat 480
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
421 aagatttgtgacttcgggtttgcacaaaattctgattccaggagatgcctacaccgattat 480
QY 481 gtagctaacgagatggtaccgagctcctgaacttcttgtgggagatactcagtatggttct 540

Db 481 gtagctaacgagatggtaccgagctcctgaacttcttgtgggagatactcagtatggttct 540
QY 541 tcagtcgatatatgggcctatttggttggcttttttcagagagctcctgacagccagccactg 600
Db 541 tcagtcgatatatgggcctatttggttggcttttttcagagagctcctgacagccagccactg 600
QY 601 tggcctggaaaaatcagatgtggaccaacttttatctgataatacagaacactagtagagacg 660
Db 601 tggcctggaaaaatcagatgtggaccaacttttatctgataatacagaacactagtagagacg 660
QY 661 gggtttcgccatgttgaccaggctggtctcgaactcttgacgtcaagtgatccacctgcc 720
Db 661 gggtttcgccatgttgaccaggctggtctcgaactcttgacgtcaagtgatccacctgcc 720
QY 721 gtacacctcaaaagtgtcgggaattacaggaaaaattaatcccaagacatcaatcaatcttt 780
Db 721 gtacacctcaaaagtgtcgggaattacaggaaaaattaatcccaagacatcaatcaatcttt 780
QY 781 aaaagtaacgggtttttccatggcatcagtcagtcacacttgagccagaagaacatggaactctt 840
Db 781 aaaagtaacgggtttttccatggcatcagtcagtcacacttgagccagaagaacatggaactctt 840
QY 841 gaggaagaattctcagatgttcacacctgtgcacctgtggtcctggaacttcacatgaagggtgtctgaag 900
Db 841 gaggaagaattctcagatgttcacacctgtgcacctgtggtcctggaacttcacatgaagggtgtctgaag 900
QY 901 atgaatccagatgacagattaaacctgttcccaactcttcccaactctggagagctcctactttgattct 960
Db 901 atgaatccagatgacagattaaacctgttcccaactcttcccaactctggagagctcctactttgattct 960
QY 961 ttccaagagcccaaaattaaaagaaaagcacgtaatgaaggaaagaaacagagacgcca 1020
Db 961 ttccaagagcccaaaattaaaagaaaagcacgtaatgaaggaaagaaacagagacgcca 1020
QY 1021 caggtacttccgctcaaaaagt 1041
Db 1021 caggtacttccgctcaaaaagt 1041

RESULT 2
AAD03813
ID AAD03813 standard; cDNA; 1068 BP.
XX
AC AAD03813;
XX
DT 19-JUN-2001 (first entry)
XX
DE Human kinase cDNA #2.
XX
KW Human; kinase; gene therapy; bioreactor; mental disorder;
KW biological disorder; ss.
XX
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 1..1068
FT /*tag= a
FT /product= "Human kinase #2"
FT /note= "The coding region does not include stop codon"
FT /partial
XX
PN WO200123579-A1.
XX
XX 05-APR-2001.
XX
XX 27-SEP-2000; 2000WO-US26621.
PF
XX 28-SEP-1999; 99US-0156511.
PR
XX (LEXI-) LEXICON GENETICS INC.
PA
XX Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
PI

XX Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
PI WPI; 2001-266166/27.
XX P-PSDB; AAE00492.
DR
XX New isolated human kinase polynucleotide useful for generating
PT antibodies, as reagents in diagnostic assays and for screening for
PT compounds useful for treating mental, biological or medical diseases
PT
XX Disclosure; Page 29-30; 38pp; English.
PS
XX The present sequence is a cDNA encoding novel human protein
CC (NHP) known as human kinase. The human kinases share structural
CC similarity with animal kinases, more particularly serine or
CC threonine protein kinases. Human kinase cDNA is useful for the
CC detection of mutant human kinase for the diagnosis of disease,
CC and also as a therapeutic. It is useful for screening drugs
CC effective in the treatment of symptomatic or phenotypic
CC manifestations perturbing the normal function of NHP in the
CC body. The NHP nucleotide sequences are useful for generation of
CC antibodies, as reagents in diagnostic assays, for the
CC identification of other cellular gene products related to human
CC kinases, and as reagents in assays for screening compounds that
CC are useful for treating mental, biological or medical disorders.
CC NHP oligonucleotides are used as probes. The labelled NHP probes
CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.
XX
SQ Sequence 972 BP; 323 A; 172 C; 214 G; 263 T; 0 other;

Query Match 79.1%; Score 823; DB 22; Length 972;
Best Local Similarity 90.6%; Pred. NO. 2e-210;
Matches 929; Conservative 0; Mismatches 0; Indels 96; Gaps 1;

QY 1 atggaaaagtatgaaaaattagctaagactggagaagggtcttatgggtgtattcaaa 60
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 61 tgcagaaaacaaacctctggacaagtagtagctgttaaaaaatttgggaatctgaagat 120
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 121 gatcctgtttaaagaaaatagcactaagagaaaatcgtatgttgaagcaattaaaaacat 180
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 181 ccaaatcttgtgaacctcatcgagggtgttcaggagaaaaaggaaaaatgcatttagtttt 240
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 241 gaatactgtgatcatacacttttaaatgagctggaaagaaacccaaatggagttgctgat 300
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 301 ggagtgatcaaaagcgtattatggcaaacacttcaagctcttaatttctgcatatacat 360
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 361 aactgtattcacagagatataaaacctgaaataattctaataagcaaggaataatc 420
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 421 aagatttgtgacttcgggtttgcacaaaattctgattccaggagatgcctacaccgattat 480
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 421 aagatttgtgacttcgggtttgcacaaaattctgattccaggagatgcctacaccgattat 480
QY 481 gtagctacgagatggtaccgagctcctgaaacttcttgtggagatactcagtagttct 540
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 541 tcagtcgatatatgggctattgttgtgtgttttgcagagctcctgacagggccagcactg 600
Db 541 tcagtcgatatatgggctattgttgtgtgttttgcagagctcctgacagggccagcactg 600
QY 601 tggcctggaaaaatcagatgtggaccacatttatctgataatcagaacact- 660
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 661 gggtttcgcatgttgaccaggctggtctcgaaactcttgacgtcaagtatccacctgcc 720
Db 651 ----- 650
QY 721 gtagcctctcaaaagtctggaattacagggaaaaattaaatcccaagacataatcaatcttt 780
Db 651 -----aggaaaaattaaatcccaagacataatcaatcttt 684
QY 781 aaaagtaacgggttttccatggcctacgtatatacctgagccagaagacatggaactctt 840
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 841 gaggaaggttctcagatgttccatcctctggctcgtgaacttcaatgaaggggtgtctgaag 900
Db 841 gaggaaggttctcagatgttccatcctctggctcgtgaacttcaatgaaggggtgtctgaag 804
QY 901 atgaatccagatgacagattaacctgttcccaactcctggagagctcctattgtattct 960
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 805 atgaatccagatgacagattaacctgttcccaactcctggagagctcctattgtattct 864
QY 961 tttcaagaggcccaaaattaaaaagaaagcacgtaatgaaggagaacagaagacgcaaa 1020
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1021 caggt 1025
Db |||||
QY 925 caggt 929

RESULT 5
AAS06725
ID AAS06725 standard; cDNA; 1083 BP.
XX
AC AAS06725;
XX
DT 12-SEP-2001 (first entry)
XX
DE Polynucleotide sequence encoding human protein kinase #25.
XX
KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
KW metabolic disorder; immune related disease; neurological disorder;
KW neurodegenerative disorder; inflammatory disorder; infectious disease;
KW reproductive disorder; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200138503-A2.
XX
PD 31-MAY-2001.
XX
PF 22-NOV-2000; 2000WO-US32085.
XX
PR 24-NOV-1999; 99US-0167482.
XX
PA (SUC-E-) SUGEN INC.
XX
PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
PI Flanagan P, Clary D;
XX

DR WPI; 2001-266166/27.
DR P-PSDB; AAE00490.
XX
PT New isolated human kinase polynucleotide useful for generating
PT antibodies, as reagents in diagnostic assays and for screening for
PT compounds useful for treating mental, biological or medical diseases
XX
PS Disclosure; Page 34-35; 38pp; English.
XX
CC The present sequence encodes a novel human protein (NHP) which
CC is a human kinase. A polymorphism was identified in the 3' UTR
CC of the present sequence. The human kinases share structural
CC similarity with animal kinases, more particularly serine or
CC threonine protein kinases. Human kinase cDNA is useful for the
CC detection of mutant human kinase for the diagnosis of disease,
CC and also as a therapeutic. It is useful for screening drugs
CC effective in the treatment of symptomatic or phenotypic
CC manifestations perturbing the normal function of NHP in the
CC body. The NHP nucleotide sequences are useful for generation of
CC antibodies, as reagents in diagnostic assays, for the
CC identification of other cellular gene products related to human
CC kinases, and as reagents in assays for screening compounds that
CC are useful for treating mental, biological or medical disorders.
CC NHP oligonucleotides are used as probes. The labelled NHP probes
CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.
XX
SQ Sequence 1819 BP; 539 A; 368 C; 406 G; 505 T; 1 other;

Query Match 68.1%; Score 709; DB 22; Length 1819;
Best Local Similarity 77.0%; Pred. No. 7.8e-180;
Matches 1025; Conservative 0; Mismatches 0; Indels 306; Gaps 1;

QY 1 atggaaaagtatgaaaaattagctaagactggagaagggtcttattggggtgtattcaaa 60
Db |||||||
QY 61 tgcagaaacaaaaccctctggacaagtagtagctgtttaaaaaatttggaaatcgaagat 120
Db |||||||
QY 198 tgcagaaacaaaaccctctggacaagtagtagctgtttaaaaaatttggaaatcgaagat 257
QY 121 gatcctgttgaagaaaaatagactaagagaaatacgtatgttgaagcaattaaacat 180
Db |||||||
QY 258 gatcctgttgaagaaaaatagactaagagaaatacgtatgttgaagcaattaaacat 317
QY 181 ccaaatcttgtgaacctcatcgaggtgttcaggagaaaaaggaaaaatgcatttagtttt 240
Db |||||||
QY 318 ccaaatcttgtgaacctcatcgaggtgttcaggagaaaaaggaaaaatgcatttagtttt 377
QY 241 gaatactgtgatcatacaccttttaaatgagctggagaaagaaaccccaatggagttgctgat 300
Db |||||||
QY 378 gaatactgtgatcatacaccttttaaatgagctggagaaagaaaccccaatggagttgctgat 437
QY 301 ggagtgatcaaaagcgtattatggaacacacttcaagctcttaatttctgtcatatacat 360
Db |||||||
QY 438 ggagtgatcaaaagcgtattatggaacacacttcaagctcttaatttctgtcatatacat 497
QY 361 aactgtattcacagagatatataaacctgaaaaataattctaataactaagcaaggaataatc 420
Db |||||||
QY 498 aactgtattcacagagatatataaacctgaaaaataattctaataactaagcaaggaataatc 557
QY 421 aagatttgtgaacttcgggttgcacaaattctg----- 453
Db |||||||
QY 558 aagatttgtgaacttcgggttgcacaaattctgagttggacttcatcttctctgtgccc 617

QY 454 ----- 453
Db 618 tccttgattggcttaataagttgaccttottgaaatttttctgccaaattcagagattttt 677
QY 454 ----- 453
Db 678 ctctggcttgatccattgctgacacagtgtttcaccatggggcccgctcatctcga 737
QY 454 ----- 453
Db 738 acttctggcctcaagtatccttccacctcggcctcccaagtgtgctggattgcaagtgtg 797
QY 454 ----- 453
Db 798 agccaccgtgccagccagatttttccaacaaataactactgagagctcacaaagattgttt 857
QY 454 ----- 474
Db 858 ttagtggaaacacaatttcgaacaaaattttgagaacgcattccaggagatgcctacacc 917
QY 475 gattatgtagctacgagatggtaccgagctcctgaacttcttgtgtgggagatactcagtat 534
Db 918 gattatgtagctacgagatggtaccgagctcctgaacttcttgtgtgggagatactcagtat 977
QY 535 ggttcttcagtcgatataatgggctatttgggtgtgtgttttgcagagctcctgacagggccag 594
Db 978 ggttcttcagtcgatataatgggctatttgggtgtgtgttttgcagagctcctgacagggccag 1037
QY 595 ccactgtggcctggaaaaatcagatgtggaccaactttatctgataatcagaaactagta 654
Db 1038 ccactgtggcctggaaaaatcagatgtggaccaactttatctgataatcagaaactagta 1097
QY 655 gagacggggtttcgccatgttgaccaggtgtgtctcgaacttctgacgtcaagtgtacca 714
Db 1098 gagacggggtttcgccatgttgaccaggtgtgtctcgaacttctgacgtcaagtgtacca 1157
QY 715 cctgcctagcctctcaaaagtgtggaatttacaggaaaaattaatcccaagacatcaatca 774
Db 1158 cctgcctagcctctcaaaagtgtggaatttacaggaaaaattaatcccaagacatcaatca 1217
QY 775 atctttaaaaagtaacgggtttttcccatggcatcagtataccctgagccagaaagacatggaa 834
Db 1218 atctttaaaaagtaacgggtttttcccatggcatcagtataccctgagccagaaagacatggaa 1277
QY 835 actcttgaggaaaaagtcttcagatgttcatctctgtggctctgaaacttcattgaagggtgtg 894
Db 1278 actcttgaggaaaaagtcttcagatgttcatctctgtggctctgaaacttcattgaagggtgtg 1337
QY 895 ctgaagatgaatccagatgacagattaaacctgttcccaactcctgagagctcctactttt 954
Db 1338 ctgaagatgaatccagatgacagattaaacctgttcccaactcctgagagctcctactttt 1397
QY 955 gattctttcaagaggcccaaaattaaaaagaaacacgtaatgaaggaaacacagaaga 1014
Db 1398 gattctttcaagaggcccaaaattaaaaagaaacacgtaatgaaggaaacacagaaga 1457
QY 1015 cgccaacacaggt 1025
Db 1458 cgccaacacaggt 1468

RESULT 7
AAF44669
ID AAF44669 standard; cDNA; 911 BP.
XX
AC AAF44669;
XX
DT 27-MAR-2001 (first entry)
XX
DE Novel protein kinase cDNA, SEQ ID NO: 49.
XX Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
KW immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic;

KW dermatological; antidiabetic; antiinfertility; gene therapy; vaccine; immune disorder; cardiovascular disease; neurodegenerative disease; cancer; autoimmune disorder; stroke; inflammatory bowel disease; inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.

OS Homo sapiens.

XX WO200073469-A2.

PN 07-DEC-2000.

XX 26-MAY-2000; 2000WO-US14842.

PF 28-MAY-1999; 99US-0136503.

XX (SUGE-) SUGEN INC.

PA Plowman GD, Martinez R, Whyte D, Sudersanam S;

PI WPI; 2001-032161/04.

XX P-PSDB; AAB65642.

DR Nucleic acids encoding kinase polypeptides, useful for diagnosing and treating immune-related diseases and disorders, cardiovascular disease, neurodegenerative diseases and/or cancers -

DR Example 4; Fig 2; 310pp; English.

XX The present sequence encodes a novel protein kinase. The nucleic acids and the protein kinases they encode may be used in the treatment and diagnosis of diseases associated with inappropriate kinase expression such as immune-related diseases and disorders, cardiovascular disease, neurodegenerative diseases and/or cancers. The nucleic acids and complementary sequences may also be used as DNA probes in diagnostic assays. The kinase polypeptides may be used as antigens in the production of antibodies of kinase expression and activity. Anti-kinase antibodies and kinase antagonists may also be used to down regulate kinase expression and activity. Diseases related to kinase expression and activity include rheumatoid arthritis, atherosclerosis, autoimmune disorders, complications of organ transplantation, myocardial infarction, immune disorders, cardiomyopathies, strokes, renal failure, oxidative-stress related disorders, chronic inflammatory bowel disease, chronic inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis, psoriasis, rhinitis, rhinitis, autoimmunity, diabetes, cancers and reproductive disorders.

XX Sequence 911 BP; 293 A; 182 C; 178 G; 258 T; 0 other;

Query Match 51.6%; Score 537; DB 22; Length 911; Best Local Similarity 87.0%; Pred. No. 5.8e-134; Matches 643; Conservative 0; Mismatches 0; Indels 96; Gaps 1;

Qy 285 aaatggagttgctgatggagtgatcaaaagcgtattatggcaaacacttcaagctcttaa 344
|||||
Db 1 aaatggagttgctgatggagtgatcaaaagcgtattatggcaaacacttcaagctcttaa 60

Qy 345 ttctgtcatatacataaactgtattcacagagatataaaacctgaaaaattcttaataac 404
|||||
Db 61 ttctgtcatatacataaactgtattcacagagatataaaacctgaaaaattcttaataac 120

Qy 405 taagcaaggaataatcaagatttgtgacttcgggtttgcacaaaattctgattccaggaga 464
|||||
Db 121 taagcaaggaataatcaagatttgtgacttcgggtttgcacaaaattctgattccaggaga 180

Qy 465 tgacctacaccgattatgtagctacgagatggtaccgagctcctgaacttcttgtgggaga 524
|||||
Db 181 tgacctacaccgattatgtagctacgagatggtaccgagctcctgaacttcttgtgggaga 240

Qy 525 tactcagtatggttcttcagtcgatatatgggctattggtgtgtttttgcagagctcct 584
|||||
Db 241 tactcagtatggttcttcagtcgatatatgggctattggtgtgtttttgcagagctcct 300

QY 585 gacagggccagccactgtggcctgggaaatcagatgtggaccacttttatctgataatcag 644
|||||
Db 301 gacagggccagccactgtggcctgggaaatcagatgtggaccacttttatctgataatcag 360

QY 645 aacactagtagagacggggtttcgccatgttgaccaggctgtctcgaaactcttgacgtc 704
|||||
Db 361 aacact----- 366

QY 705 aagtgatccacctgccgtagcctctcaaaagtctggaattacaggaaaaattaatcccaag 764
|||||
Db 367 -----aggaaaaattaatcccaag 384

QY 765 acatcaatcaatcttttaaaagtaacgggtttttccatggcatcagatacctgagccaga 824
|||||
Db 385 acatcaatcaatcttttaaaagtaacgggtttttccatggcatcagatacctgagccaga 444

QY 825 agacatggaaaactcttgaggaaaagtcttcagatgttcacatgttggtcttgaacttcat 884
|||||
Db 445 agacatggaaaactcttgaggaaaagtcttcagatgttcacatgttggtcttgaacttcat 504

QY 885 gaaggggtgtctgaagatgaatccagatgacagattaacctgttcccaactcctggagag 944
|||||
Db 505 gaaggggtgtctgaagatgaatccagatgacagattaacctgttcccaactcctggagag 564

QY 945 ctctactttgattcttttcaagaggccccaattaaagaaaaagcagtaatgaaggaag 1004
|||||
Db 565 ctctactttgattcttttcaagaggccccaattaaagaaaaagcagtaatgaaggaag 624

QY 1005 aaacagaagacgcccaacag 1023
|||||

Db 625 aaacagaagacgcccaacag 643

RESULT 8

AAD03812

ID AAD03812 standard; cDNA; 561 BP.

XX

AC AAD03812;

XX 19-JUN-2001 (first entry)

XX Human kinase cDNA #1.

DE Human; kinase; gene therapy; bioreactor; mental disorder;

XX Human; kinase; gene therapy; bioreactor; mental disorder;

KW biological disorder; ss.

XX Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 1..561

FT /*tag= a

FT /product= "Human kinase #1"

FT /note= "The coding region does not include stop codon"

FT /partial

XX WO200123579-A1.

XX

XX 05-APR-2001.

PD

PF 27-SEP-2000; 2000WO-US26621.

XX

PR 28-SEP-1999; 99US-0156511.

XX

PA (LEXI-) LEXICON GENETICS INC.

XX

PI Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;

XX WPI; 2001-266166/27.

DR P-PSDB; AAE00490.

DR

XX New isolated human kinase polynucleotide useful for generating for antibodies, as reagents in diagnostic assays and for screening for compounds useful for treating mental, biological or medical diseases -

XX Disclosure; Page 27; 38pp; English.

PS The present sequence is a cDNA encoding novel human protein

XX (NHP) known as human kinase. The human kinases share structural

CC similarity with animal kinases, more particularly serine or

CC threonine protein kinases. Human kinase cDNA is useful for the

CC detection of mutant human kinase for the diagnosis of disease,

CC and also as a therapeutic. It is useful for screening drugs

CC effective in the treatment of symptomatic or phenotypic

CC manifestations perturbing the normal function of NHP in the

CC body. The NHP nucleotide sequences are useful for generation of

CC antibodies, as reagents in diagnostic assays, for the

CC identification of other cellular gene products related to human

CC kinases, and as reagents in assays for screening compounds that

CC are useful for treating mental, biological or medical disorders.

CC NHP oligonucleotides are used as probes. The labelled NHP probes

CC are useful for screening human genomic library for identifying

CC polymorphisms and as primers in amplification assays to detect

CC mutations within the exons, introns and splice sites that can

CC be used in diagnostics and pharmacogenomics. Nucleotide construct

CC encoding NHP products are used to genetically engineer cells

CC in vivo that functions as bioreactors in the body delivering a

CC continuous supply of NHP to the body. Nucleotide constructs

CC encoding functional NHPs are used in gene therapy for the

CC modulation of NHP expression.

XX Sequence 561 BP; 189 A; 87 C; 112 G; 173 T; 0 other;

SQ

Query Match 43.7%; Score 454.4; DB 22; Length 561;

Best Local Similarity 99.8%; Pred. No. 5.8e-112;

Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 atggaaaagtatgaaaaattagctaagactggagagggtcttattatgggtgtattcaaa 60

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 61 tgcagaaacaaaacctctggacaagtagtagctgtttaaaaaatttgggaatctgaagat 120

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 121 gatcctgtgttaagaaaaatagcactaagagaaatacgtatgttgaagcaattaaaaacat 180

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 181 ccaaatcttgtgaacctcatcgagggtgttcaggagaaaaaggaaaatgcatttagtttt 240

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 241 gaatactgtgatcatcacacttttaaatgagctggaaagaaaccccaaatggagtgtgat 300

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 301 ggagtgatcaaaagcgttattatggcaaacacttcaagctcttaatttctgtcatatacat 360

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 361 aactgtattcacagagatataaaacctgaaaaatttctaataactaagcaagggaataatc 420

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 421 aagatttgtgacttcgggtttgcacaaaattctgatt 456

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 421 aagatttgtgacttcgggtttgcacaaaattctgagt 456

RESULT 9

AAD03815

ID AAD03815 standard; cDNA; 594 BP.

XX

AC AAD03815;

XX

DT 19-JUN-2001 (first entry)

XX Human kinase cDNA #4.

DE

XX Human; kinase; gene therapy; bioreactor; mental disorder;

KW biological disorder; ss.

KW

XX Homo sapiens.

OS

XX

FH Key Location/Qualifiers

DE 1..594

FT /*tag= a

FT /product= "Human kinase #4"

FT /note= "The coding region does not include stop codon"

FT /partial

XX

PN WO200123579-A1.

XX

PD 05-APR-2001.

XX

PF 27-SEP-2000; 2000WO-US26621.

XX

PR 28-SEP-1999; 99US-0156511.

XX

PA (LEXI-) LEXICON GENETICS INC.

XX

PI Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;

XX

DR WPI; 2001-266166/27.

DR P-PSDB; AAE00493.

XX

PT New isolated human kinase polynucleotide useful for generating

PT antibodies, as reagents in diagnostic assays and for screening for

PT compounds useful for treating mental, biological or medical diseases -

XX

PS Disclosure; Page 31; 38pp; English.

XX

CC The present sequence is a cDNA encoding novel human protein

CC (NHP) known as human kinase. The human kinases share structural

CC similarity with animal kinases, more particularly serine or

CC threonine protein kinases. Human kinase cDNA is useful for the

CC detection of mutant human kinase for the diagnosis of disease,

CC and also as a therapeutic. It is useful for screening drugs

CC effective in the treatment of symptomatic or phenotypic

CC manifestations perturbing the normal function of NHP in the

CC body. The NHP nucleotide sequences are useful for generation of

CC antibodies, as reagents in diagnostic assays, for the

CC identification of other cellular gene products related to human

CC kinases, and as reagents in assays for screening compounds that

CC are useful for treating mental, biological or medical disorders.

CC NHP oligonucleotides are used as probes. The labelled NHP probes

CC are useful for screening human genomic library for identifying

CC polymorphisms and as primers in amplification assays to detect

CC mutations within the exons, introns and splice sites that can

CC be used in diagnostics and pharmacogenomics. Nucleotide construct

CC encoding NHP products are used to genetically engineer cells

CC in vivo that functions as bioreactors in the body delivering a

CC continuous supply of NHP to the body. Nucleotide constructs

CC encoding functional NHPs are used in gene therapy for the

CC modulation of NHP expression.

XX

SQ Sequence 594 BP; 203 A; 94 C; 115 G; 182 T; 0 other;

Query Match 43.7%; Score 454.4; DB 22; Length 594;

Best Local Similarity 99.8%; Pred. No. 6e-112;

Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 atggaaaaagtatgaaaaaattagctaagactggagagggtcttattgggtgtattcaaa 60

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 61 tgcagaaacaaaacctctctggacaaagtagtagctgttaaaaaatttgtggaatctgaagat 120

Db 61 tgcagaaacaaacctctggacaagtagctgttaaaaaattgtggaatctgaagat 120
QY 121 gatcctgttgaagaaatagcactaagagaaatacgtatottgaagcaattaaaacat 180
Db 121 gatcctgttgaagaaatagcactaagagaaatacgtatgttgaagcaattaaaacat 180
QY 181 ccaaatcttgaacctcatcgagggtgttcaggagaaaaaggaaaaatgcatttagtttt 240
Db 181 ccaaatcttgaacctcatcgagggtgttcaggagaaaaaggaaaaatgcatttagtttt 240
QY 241 gaatactgtgatcatacacttttaaatgagctggaaagaaaccccaaatggagttgctgat 300
Db 241 gaatactgtgatcatacacttttaaatgagctggaaagaaaccccaaatggagttgctgat 300
QY 301 ggagtgatcaaaaagcgtattatggcaaacacttcaagctctLaattttctgtcatatacat 360
Db 301 ggagtgatcaaaaagcgtattatggcaaacacttcaagctcttaattttctgtcatatacat 360
QY 361 aactgtattcacagagatataaaacctgaaaatattcttaactaaactaagcaaggaataatc 420
Db 361 aactgtattcacagagatataaaacctgaaaatattcttaactaaactaagcaaggaataatc 420
QY 421 aagattttgacttcgggtttgcacaaaattctgatt 456
Db 421 aagattttgacttcgggtttgcacaaaattctgagt 456

RESULT 10
AAF44670
ID AAF44670 standard; cDNA; 2615 BP.
XX
AC AAF44670;
XX
DT 27-MAR-2001 (first entry)
XX
DE Novel protein kinase cDNA, SEQ ID NO: 50.
XX
KW Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
KW immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic;
KW dermatological; antidiabetic; antiinferility; gene therapy; vaccine;
KW immune disorder; cardiovascular disease; neurodegenerative disease;
KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
KW inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.
OS Mus musculus.
XX
XX WO200073469-A2.
PN
XX
PD 07-DEC-2000.
XX
XX 26-MAY-2000; 2000WO-US14842.
PF
XX
PR 28-MAY-1999; 99US-0136503.
XX
PA (SUGE-) SUGEN INC.
XX
XX Plowman GD, Martinez R, Whyte D, Sudersanam S;
PI
XX WPI; 2001-032161/04.
DR P-PSDB; AAB65643.
XX

PT Nucleic acids encoding kinase polypeptides, useful for diagnosing and
PT treating immune-related diseases and disorders, cardiovascular disease,
PT neurodegenerative diseases and/or cancers -
XX
PS Disclosure; Fig 2; 310pp; English.

CC The present sequence encodes a novel protein kinase. The nucleic acids
CC and the protein kinases they encode may be used in the treatment and
CC diagnosis of diseases associated with inappropriate kinase expression
CC such as immune-related diseases and disorders, cardiovascular disease,
CC neurodegenerative diseases and/or cancers. The nucleic acids and

CC complementary sequences may also be used as DNA probes in diagnostic
CC assays. The kinase polypeptides may be used as antigens in the production
CC of antibodies of kinase expression and activity. Anti-kinase antibodies
CC and kinase antagonists may also be used to down regulate kinase
CC expression and activity. Diseases related to kinase expression and
CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
CC disorders, complications of organ transplantation, myocardial infarction,
CC immune disorders, cardiomyopathies, strokes, renal failure,
CC oxidative-stress related disorders, chronic inflammatory bowel disease,
CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
CC reproductive disorders.
XX

SQ Sequence 2615 BP; 742 A; 580 C; 585 G; 708 T; 0 other;

Query Match 43.5%; Score 453.2; DB 22; Length 2615;
Best Local Similarity 73.1%; Pred. No. 2.2e-11;
Matches 658; Conservative 0; Mismatches 143; Indels 99; Gaps 2;

QY 135 gaaaaatagcactaagagaaaatcgtatgttgaagcaattaaaacatccaaatcttgtgaa 194
Db 1 gaaaaatagccctgcgggaaatccgtatgctgaag--ttgaaacaccccaaacctcgtgaa 57
QY 195 cctcatcgagggtgttcaggagaaaaaggaaaatgcatttagttttgaataactgtgatca 254
Db 58 cctcatcgagggtgttcagaagaaagagaagatgcatactagtttttgagtactgtgatca 117
QY 255 tacacttttaaatgagctgggaaagaaaccccaaatggagttgctgatggagtgatcaaaaag 314
Db 118 cacactgttaacgagctggagagaaaccccaaacgagtttctgatggagtgattaaaag 177
QY 315 cgtattatggcacaacacttcaagctctttaatttctgtcatatacatataactgtattcacag 374
Db 178 tgtgctatggcaaaccttcaagcccttaacttctgtcacaaagcacaattgtattcatcg 237
QY 375 agataataaacctgaaaaatattctaataactaagcaaggaaataatcaagattttgtgactt 434
Db 238 ggatgtaaacacctgaaaaacatcctaataaccaagcaagggatgataaagattttgtgactt 297
QY 435 cgggttgcacaaaattctgattccaggagatgcctacaccgattatgtagctacgagatg 494
Db 298 tggatttgcacgaattcttaattccaggagacgcctcacagactatgttgcaccagggtg 357
QY 495 gtaccgagctcctgaaacttcttgtgggagatactcagatggtttcagtcgatatatg 554
Db 358 gtaccgagcccccgaaacttctctgtgggagacacgaagtacggttctcctctgtagacgtgtg 417
QY 555 ggctattgtgtgttttgcagagctcctgacagccagccactgtggcctggaaaaatc 614
Db 418 ggccgtcggctgtgttttgcagagctcctgacgggtcagccactctggccgggaaaaatc 477
QY 615 agatgtggaccaactttatctgataatcagaacacactagtagagacggggttctgcctatgt 674
Db 478 cgacgtggaccagctttacctgatcatcaggaacttg----- 514
QY 675 tgaccaggctggtctcgaaactcttgaagtcgaagtcatccactgccgtagcctctcaaaag 734
Db 515 ----- 514
QY 735 tgcgtggaattacaggaaaaattaatcccaagacatcaatcaatctttaaaagtaacgggtt 794
Db 515 -----gggaagctgattccaagacacacagtcatactttaggagtaaccagtt 561
QY 795 ttcccatggcatcagtatcacctgagccagaagacatggaaaactcttgagggaaaagtcttc 854
Db 562 ttcccgggcatcagcatcacctgaaaccagaggacatggagactcttgaagaaaaattcttc 621
QY 855 agatgttcatcctgtggtctgaaacttcaatgaaggggtgtctgaagatgaatccagatga 914
Db 622 aaatgttcagcctgtggtcttaagtttcatgaaggggatgctgaagatgaatcctgatga 681
QY 915 cagattaaacctgttcccaactcctggagagctcctactcttatttccaagaggccca 974

Db	682	gaggctgacctgtgccagctgctggacagtgacctactttgagtcctttcaagaggatca	741
QY	975	aattaaagaaagcacgtaatatgaagaaagaaacagaaagacgccaacaggtacttcogct	1034
Db	742	aatgaaaaagaaagcccgagtgaggggagaaagcgcgaagcgccagcagaatcaactgct	801
RESULT 11			
AAA29746			
ID	AAA29746	standard; DNA; 1698 BP.	
XX	AC	AAA29746;	
XX	DT	15-AUG-2000 (first entry)	
XX	DE	Rabbit KKIAMRE kinase genomic DNA SEQ ID NO:5.	
XX	DE	Rabbit KKIAMRE kinase; learning-induced kinase; learning; memory;	
KW	KW	Rabbit; KKIAMRE kinase; learning-induced kinase; learning; memory;	
KW	KW	cdc2-related kinase; brain; gene therapy; genetic disorder; detection;	
KW	KW	identification; ds.	
XX	OS	Oryctolagus cuniculus.	
XX	XX	WO200020567-A2.	
PN	PN	13-APR-2000.	
XX	PF	01-OCT-1999; 99WO-US23010.	
XX	XX	02-OCT-1998; 98US-0102906.	
PR	PR	(UYSC-) UNIV SOUTHERN CALIFORNIA.	
XX	PA	Thompson RE, Gomi H, Sun W;	
XX	PI	WPI; 2000-328932/28.	
XX	PI	P-PSDB; AAY90724.	
XX	DR		
XX	XX		
XX	XX		
PT	PT	Novel learning induced kinase polynucleotides and polypeptides, useful	
PT	PT	for the analysis of learning and memory, and for gene therapy -	
XX	XX	Claim 1; Fig 9; 64pp; English.	
PS	PS	The present sequence encodes a learning-induced kinase, designated	
XX	CC	KKIAMRE kinase, which is isolated from rabbit brain tissue. KKIAMRE	
CC	CC	kinase is a cdc2-related kinase. The KKIAMRE kinase polynucleotides can	
CC	CC	be used to express recombinant protein for analysis, characterisation or	
CC	CC	therapeutic use, as markers for tissues in which the protein is	
CC	CC	preferentially expressed, as molecular weight markers on Southern gels,	
CC	CC	as chromosome markers or tags, to compare endogenous DNA sequences in	
CC	CC	patients to identify potential genetic disorders, as probes to hybridise	
CC	CC	and discover novel related sequences, as a source of PCR primers, and as	
CC	CC	an antigen to induce anti-DNA antibodies. The polypeptides can be used	
CC	CC	in assay to discover biological activity, to raise antibodies, as tissue	
CC	CC	markers, and to isolate correlative receptors or ligands. The	
CC	CC	polynucleotides may also be used for gene therapy for the treatment of	
CC	CC	disorders which are mediated by KKIAMRE kinase.	
XX	XX	Sequence 1698 BP; 551 A; 311 C; 390 G; 446 T; 0 other;	
SQ	SQ		
Query Match 24.2%; Score 251.8; DB 21; Length 1698;			
Best Local Similarity 62.9%; Pred. No. 1.3e-57;			
Matches 407; Conservative 0; Mismatches 237; Indels 3; Gaps 1;			
QY	1	atggaaaaagtatgaaaaattagctaagactggagaaaggggtctttatgggggtgtattcaaa	60
Db	1	atggaaaaaatatgagaaccttggttggttggaagagggttatggaatggtgatgaag	60
QY	61	tgcagaaacaaaacctcttggaagaagtagctgtttaaaaaattgtggaatctgaagat	120
Db	61	tgtaggataaagatagtgggaagaattgtggccatcaagaagttcctctagaaagtgtgat	120

QY	121	gatactgttgttaagaaaaatagcactaagagaaaaatacgtatgttgaagcaattaaaaacat	180
Db	121	gacaaaaatggttaaaaaaattgctatcgagagaaatacaagttactaaagcaactgaggcat	180
QY	181	ccaaatcttgtgaacctcactcaggtgttcaggagaaaaaaggaataatgcatttagttttt	240
Db	181	gaaaatttgtgaatctgttggaggtgtgtataaaaaaaaacgattggtacctagtcttt	240
QY	241	gaatactgtgatcatcacacttttaaatgagctggagaaagaaacccaaatggagtgtgat	300
Db	241	gaatttgtgaccacacgattcttgatgacttggaaactcttccaaatggactagatgac	300
QY	301	ggagtgatcaaaaagcgtattatggcaaacacattcaagctcttaatttctgtcatatacat	360
Db	301	caagtagtlcaaaaagtatttgtttcagattattaaatggaattggtttgtcacagtac	360
QY	361	aactgtattccacagagatatataaacctgaaaaatatattctaataactaaaggaataatc	420
Db	361	aatatcatatagagatatataaagccagagaaatatattggtctccagctcgtgcgttgtc	420
QY	421	aagatttgtgacttcgggtttgcacaaaattctg---attccaggagatgcctacaccgat	477
Db	421	aagttatgtgattttggatttgcacggacactggcagctcccggagaggtttacactgat	480
QY	478	tatgtagtacgagatggttacccgagctcctgaacttcttgtgtggagatactcagtatggt	537
Db	481	tatgtggcaactcgtggttacagagctccagaaactactgtgtgtgatgtcaagtatggc	540
QY	538	tcttcagtcgatatatatgggttatgtgtgtgttttgcagagctcctgcagagccagcca	597
Db	541	aaagctgtggatgtgtgggccattggtgtctgtgtaactgaaatgctcatggtgggaaccc	600
QY	598	ctgtggcctggaaaaatcacagatgtggaccaactttatctgtataaatcag	644
Db	601	ctgttctcgtgagactctgatattgatcagctttatcttattatgag	647
RESULT 12			
AAA29745			
ID	AAA29745	standard; cDNA; 3080 BP.	
XX	AC	AAA29745;	
XX	DT	15-AUG-2000 (first entry)	
XX	DE	Rabbit KKIAMRE kinase encoding cDNA SEQ ID NO:3.	
XX	KW	Rabbit; KKIAMRE kinase; learning-induced kinase; learning; memory;	
KW	KW	cdc2-related kinase; brain; gene therapy; genetic disorder; detection;	
KW	KW	identification; ss.	
XX	OS	Oryctolagus cuniculus.	
XX	FH	Key Location/Qualifiers	
FT	CDS	249..1949	
FT		/*tag= a	
FT		/product= "KKIAMRE kinase"	
XX	XX	WO200020567-A2.	
PN	PN	13-APR-2000.	
XX	XX	01-OCT-1999; 99WO-US23010.	
XX	XX	02-OCT-1998; 98US-0102906.	
XX	XX	(UYSC-) UNIV SOUTHERN CALIFORNIA.	
XX	XX	Thompson RE, Gomi H, Sun W;	
XX	XX	WPI; 2000-328932/28.	
XX	XX	P-PSDB; AAY90724.	

[illegible]

RESULT 14
AAS06724
ID AAS06724 standard; cDNA; 1773 BP.
XX
AC AAS06724;
XX
DT 12-SEP-2001 (first entry)
XX
DE polynucleotide sequence encoding human protein kinase #24.
XX
KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
KW metabolic disorder; immune related disease; neurological disorder;
KW neurodegenerative disorder; inflammatory disorder; infectious disease;
KW reproductive disorder; gene therapy; ss.

OS	Homo sapiens.		
XX			
PN	WO200138503-A2.		
XX			
PD	31-MAY-2001.		
XX			
PF	22-NOV-2000; 2000WO-US32085.		
XX			
PR	24-NOV-1999; 99US-0167482.		
XX			
PA	(SUGE-) SUGEN INC.		
XX			
PI	Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;		
PI	Flanagan P, Clary D;		
XX			
DR	WPI; 2001-343950/36.		
DR	P-PSDB; AAUQ3524.		

XX Nucleic acids encoding human kinase polypeptides, useful for preventing
PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT several associated diseases, and microbial infections -

XX
PS
Example 1: Figure 1: 433pp: English.

XX
GC
ACCGCTGAT-AGCAGTGGT encode for novel human protein kinases #1-57. The

novel protein kinases have been identified as members of the cytochrome or serine/threonine kinase (PTK and STK) families. The polynucleotides encoding protein kinases and the polypeptides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate kinase expression. For example, they may be used to treat inappropriate kinase expression. For example, they may be used to treat cancers (especially cancers of haematopoietic origin), cardiovascular disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes), immune related diseases (e.g. rheumatoid arthritis), neurological disorders (e.g. schizophrenia), neurodegenerative disorders (e.g. Parkinson's disease), inflammatory disorders (e.g. asthma), infectious disease (e.g. HIV) and reproductive disorders (e.g. infertility). Additionally, polynucleotides encoding protein kinases may be used for gene therapy and as DNA probes in diagnostic assays. The protein kinase polypeptides may be used as antigens in the production of antibodies against the protein kinases and in assays to identify modulators of protein kinase expression and activity.

XX
SQ
Sequence 1773 BP; 650 A; 300 C; 362 G; 461 T; 0 other;

Query Match	19.5%;	Score 202.6;	DB 22;	Length 1773;
Best Local Similarity	59.4%;	Pred. No. 1.9e-44;		
Matches 381;	Conservative	0;	Mismatches 254;	Indels 6;
			Gaps	2;

QY	1	atggaaaaagtatgaaaaattagctaagactggagaagggtcttatgggggtgtgattcaaa	60
Db	1	atggagatgtatgaacccttgaaaagtgggagaggaagttaaggaaacagtoatgaaa	60
QY	61	tgcagaaaaaacctctggacaagttagctgttaaaaaaatttgtggaatctgaagat	120
Db	61	tgtaaacataagaatactgggcagatagtgccattagaatatatttatgagagaccagaa	120
QY	121	gacctgtgtttaagaaaaatagcactaaagagaaatacgtatgttgaaagcaattaaaaacat	180
Db	121	caatctgtc---aacaaaaattgcgatgagagaaataaagtttctaaaggcaatttcacac	177
QY	181	ccaaatcttgtgaacctcatcgaggtgttcaggagagaaaaaggaaaaatgcatttagttttt	240
Db	178	gaaaacctgtcaactctgattgaagtttttagacagaaaaagaaaattcatttgggtattt	237
QY	241	gaatactgtgatacatcacacttttaaatgagctggaaaagaaaccccaaatggagttgctgat	300
Db	238	gaatttattgaccacagtattagatgagttacaacattattgtcatggactagagagt	297
QY	301	ggagtgatcaaaagcgtattattggcaaacacattcaagctcttaatttctgtcatatacat	360
Db	298	aagcgacttagaaaaatacctcttcagatccttcgagcaattgactatcttcacagtaat	357
QY	361	aactgtattcacagagatatataaacctgaaaaataattctataactaagcaaggaaataatc	420
Db	358	aatatcatcatcgagatatataaacctgagaataattttagtattcccgatcaggaaattact	417
QY	421	aagatttgtgacttcgggtttgcacaaaattct---gattccaggagatgcctacacccgat	477
Db	418	aagctctgtgatttttggttttgcagaaacactagcagctcctgggacatttatacggac	477
QY	478	tatgtagctaacagatggttacccgagctcctgaactcttgtgggagatactcagtatggt	537
Db	478	tatgtggccacacgctggtatagagctcccgaaatagattattaaagatacttcttatgga	537
QY	538	tottcagtcgatataatgggctatttggttgtgttttttgcagagctcctgacaggccagcca	597
Db	538	aaacctgtggatatctgggctttgggctgtatgatcattgagatggccactggaaatccc	597
QY	598	ctgtggccctgaaaaatcagatgtggaccacaatttatctgat	638
Db	598	atatcttcctaatagttctgaatttggaatttgaatttactccataaaat	638

RESULT 15

AAZ51207

AAZ51207
ID AAZ51207 standard: cDNA: 1667 BP.

XX
XX

AC AA251207;
XX
DT 06-JUN-2000 (first entry)
XX
DE Rat lost in leukaemia kinase beta isoform (LLK beta) cDNA.
XX
DE
XX
KW Lost in leukaemia kinase; LLK beta; rat; cdc2-related kinase; leukaemia;
KW MAPK; mitogen activated protein kinase; tumour suppressor; diagnosis;
KW cytostatic; mitotic index; treatment; prevention; gene therapy; ss.
XX
OS Rattus sp.
XX
XX
FH Key Location/Qualifiers
FT CDS 64..1437
FT /*tag= a
FT /product= "Rat lost in leukaemia kinase beta protein"
XX
PN WO200012719-A1.
XX
XX
PD 09-MAR-2000.
XX
XX
PF 31-AUG-1999; 99WO-CA00794.
XX
PR 31-AUG-1998; 98CA-2243784.
PR 20-NOV-1998; 98CA-2251249.
XX
PA (ONTA-) ONTARIO CANCER INST.
XX
PI Zanke B, Haq R, Randall S, Midmer M;
XX
DR WPI; 2000-237880/20.
DR P-PSDB; AAY70125.
XX
PT Isolated polynucleotide encoding a lost in leukemia kinase (LLK)
PT protein, useful for treatment, diagnosis and prevention of leukemia -
XX
PS Claim 4b; Page 48; 69pp; English.
XX

The present sequence is the cDNA encoding the beta isoform of rat lost in leukaemia kinase (LLK beta), isolated from rat jejunum cDNA library. It is expressed strongly in muscle, heart, liver, brain and in tissues which have a very low mitotic index. Rat LLK beta protein is closely related to cdc2-related kinases, that are putative tumour suppressors and to mitogen activated protein kinases (MAPKs). This sequence has cytotstatic activity. It is useful for the treatment, diagnosis and prevention of acute leukaemia and is also used in gene therapy.

Sequence 1667 Bp; 525 A; 353 C; 392 G; 397 T; 0 other;

Query Match 18.3%; Score 191; DB 21; Length 1667;
Best Local Similarity 57.9%; Pred. No. 2.3e-41;
Matches 379; Conservative 0; Mismatches 270; Indels 6; Gaps 2;
QY 1 atggaaaagtatgaaaaattagctaagactggagaagggtcttattggtgtattcaaa 60
Db 64 atggagatgtatgaaaccccttgaaaaagtaggagagggttatggaaacggtcatgaa 123
QY 61 tgcagaaacaaaaacctctggacaagtagtagctgtgttaaaaaatttggaaatctgaagat 120
Db 124 tgcaagcataaggatactggcggtatgtggccattgaagatatattctatgagaaaccagaa 183
QY 121 gatcctgtgtttaagaaaaatagcactaagagaaatacgtatgttgaagcaattaaaacat 180
Db 184 aaatctgtc---aacaataattgcaacgagagaaataaagtttctaaagcaatttcgtcat 240
QY 181 ccaaatcttgtgaacctcatcgagggtgttcaggagagaaaaaggaaaaatgcattagtttt 240
Db 241 gaaaacctgggtcaatctgtgattgaagtttttagacaaaaaaagaaatccatttgggtatt 300
QY 241 gaatactgtgatcatcacacttttaaatgagctggaaagaaaccccaaatgagttgctgat 300
Db 301 gagtttattgaccacacggtcttagatgagttacagcattattgtcagggattagagagt 360

QY 301 ggagtgatcaaaaagcgtattattatggcaaacacacttcaagctcttaatttctgtcatatacat 360
Db 361 aagcgactaaagaaagtacotgtttccagatctccgagcgatcgagtatctgcacaacaac 420
QY 361 aactgtattcacagagagatatataaaacotgaaaaatattcttaataactaaagcaagggaataatc 420
Db 421 aaattatccaccgagatatataaagcctgagaatatttttagtctcccagtcaggaattaca 480
QY 421 aagatttgtgacttcgggttttgcaaaaattct---gattccaggagatgcctacacgat 477
Db 481 aagctctgcgatttttggttttgcgcggacactagcagctcctggagacgttttacacagac 540
QY 478 tatgtagctacgagatggttacccgagctcctggaacttcttggggagagatactcagttatggt 537
Db 541 tacgtggccaccccgctggttacagagctccagagctggtgttgaaagacacaccacctatgga 600
QY 538 tcttcagtcgatataatgggctattggtgtgttttttgcagagctcctgacagggccagcca 597
Db 601 aaaccagtggtatctctgggttgggtgtgtatgatcattgaaatggccactggcaatccc 660
QY 598 ctgtggcctggaaaaatcagatgtggaccaacttttatctgalaatcagaacacactag 652
Db 661 taccttctcagcagttcccgatttggatttggctccacaagattgttttaaaagtag 715

Search completed: May 2, 2002, 17:13:01
Job time: 5438 sec

.....

Fri May 3 11:02:13 2002

us-09-671-050-9.rng

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 14:02:28 ; Search time 2235.62 Seconds
(without alignments)
5003.693 Million cell updates/sec

Title: US-09-671-050-9
Perfect score: 1041
Sequence: 1 atggaagaagtatgaaaaatt.....aggtacttcgcgtcaaaaagt 1041

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST: *
1: em_estfun: *
2: em_esthum: *
3: em_estin: *
4: em_estom: *
5: em_estpl: *
6: em_estba: *
7: em_estro: *
8: em_estov: *
9: em_htc: *
10: gb_estl: *
11: gb_est2: *
12: gb_htc: *
13: gb_gss: *
14: em_gss_fun: *
15: em_gss_hum: *
16: em_gss_inv: *
17: em_gss_pln: *
18: em_gss_pro: *
19: em_gss_rod: *
20: em_gss_vrt: *
21: em_gss_other: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	333.6	32.0	500	10	AI385966 ml34h09.y
2	331	31.8	1691	12	AK016781 Mus muscu
3	316.8	30.4	938	10	AA061797 ml34h09.r
4	258.6	24.8	731	10	AW106692 um32d03.y
5	255	24.5	718	10	AA286088 vc33b05.r
6	209.8	20.2	579	10	AW233105 fj28b10.y
7	199	19.1	295	10	AA626859 zu89f10.s
8	186.6	17.9	727	10	AI508835 vc33b05.y
9	175.6	16.9	1535	12	BC010966 Homo sapi
10	174.8	16.8	582	10	AV672492 AV672492
11	174.4	16.8	665	13	AQ953706 Sheared D
12	158.4	15.2	495	10	AI385990 ml150h06.y

13	154.6	14.9	744	11	BF787339	BF787339	602113731
14	152.6	14.7	879	11	BG281841	BG281841	602403067
15	152.6	14.7	926	10	AL524036	AL524036	AL524036
16	150.4	14.4	779	11	BG760621	BG760621	602717373
17	149.6	14.4	384	10	BE223187	BE223187	kp73g08.y
18	141.4	13.6	808	10	AI746896	AI746896	u108g10.y
19	141.2	13.6	795	10	BE275968	BE275968	601120739
20	140	13.4	957	11	BI410031	BI410031	602962512
21	135.2	13.0	699	10	AL584951	AL584951	AL584951
22	134.8	12.9	754	11	BG757858	BG757858	602711380
23	134.6	12.9	1131	10	BE563495	BE563495	601334976
24	134	12.9	202	10	BE135229	BE135229	ug28e10.y
25	134	12.9	696	11	BF479244	BF479244	L48-2906T
26	133.8	12.9	752	10	BE746563	BE746563	601580135
27	133.8	12.9	911	11	BG397188	BG397188	602434149
28	133	12.8	680	10	AL585015	AL585015	AL585015
29	133	12.8	1112	10	AT002436	AT002436	AT002436
30	132.8	12.8	619	11	BE844514	BE844514	EST259 Ap
31	132.4	12.7	764	11	BG162489	BG162489	df15e08.y
32	132.4	12.7	1000	11	BF316057	BF316057	601896036
33	131.2	12.6	831	11	BG825471	BG825471	602749632
34	131	12.6	297	10	AI965145	AI965145	f87e07.y
35	131	12.6	963	11	BG420755	BG420755	602448546
36	130.4	12.5	613	10	AL504123	AL504123	AL504123
37	130.4	12.5	778	10	BE413170	BE413170	MCG014.CO
38	130.2	12.5	793	11	BG700342	BG700342	602680030
39	130.2	12.5	796	11	BG915878	BG915878	602815329
40	130	12.5	713	10	AI246661	AI246661	q127c01.x
41	129.8	12.5	379	10	AI694352	AI694352	wd45g08.x
42	129.8	12.5	884	11	BG440417	BG440417	GA_Ea000
43	129.8	12.5	945	11	BG481522	BG481522	602528424
44	129.6	12.4	673	11	BG976698	BG976698	602847026
45	129.2	12.4	732	10	AW701949	AW701949	uq93b12.y

ALIGNMENTS

RESULT 1
AI385966 500 bp mRNA EST 27-JAN-1999
LOCUS ml34h09.y1 Stratagene mouse testis (#937308) Mus musculus cDNA
DEFINITION clone IMAGE:513953 5', similar to gb:X66358 SERINE/THREONINE-PROTEIN KINASE KKIALRE (HUMAN);, mRNA sequence.
ACCESSION AI385966 GI:4199429
VERSION AI385966.1
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 500)
AUTHORS Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE The WashU-NCI Mouse EST Project 1999
JOURNAL Unpublished (1999)
COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:307801
This read is a RESEQUENCE of a previously sequenced mouse clone
This read has been verified (found to hit its original self in the correct orientation)
Seq primer: -40RP from Gibco
High quality sequence stop: 339.

Db	358	GTACCGAGCCCCGAAACTTCGCTGGGAGACACGAAGTACGGTTCCTCTCTGAGACGTGTG	417
Qy	555	ggctattggtgtttttgcagagctcctgacagccagccactgtggcctggaataac	614
Db	418	GGCCGTCGGCTGTGTTTTCGAGA-CTCCTGACGGGTGAG-CACTCTGGCCGGGAAACC	475
RESULT	4		
AW106692			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
BASE COUNT			
ORIGIN			
Query Match			
Best Local Similarity			
Matches			
QY	5	aaaagtatgaaaaattagctaagactggagagggtcttatgggtgtgtattccaatgca	64

Db	156	AAATATGAAAAAAATTTGGAAAGATTGGAGAAAGGCTCTATGGGGTAGTGTTCAAAGTGCA	215
QY	65	gaaacaaaaacctctgtggacaagtagtagtctgtttaaaaaatttggaaatctgaagatgac	124
Db	216	GAACACAGGGACACGGGTCAGATCGTGGCCATCAAGAGGTTTCTGGAAACCGAAGATGACC	275
QY	125	ctgtgtttaagaaaaatagcactaagagaaatacgtatgttgaagcaattaaaaacatccaa	184
Db	276	CTGTCATAAAGAAAAATCGCCCTTCGAGAAATCCGCATGCTCAAGCAACTCAAGCACCCCA	335
QY	185	atcttgtgaacctcatcgaggtgttcaggagagaaaaatgcatttagtttttgaat	244
Db	336	ACCTGGTCAACCTCCTGGAAGTCTTCCGGAGGAAGCGGAGGCTTCACCTGGTGTTCGAGT	395
QY	245	acttgtatcatcacacttttaaatgagctggaagaaacccccaaatggagttgctgatggag	304
Db	396	ACTGCGACCACACGGTGTCTCACGAGCTGGATCGGTATCAGAGGGGGGTACCAGAGCCTC	455
QY	305	tgatcaaaagcgtattattatggcacaacaccttcaagctcttaatttctgtcatatacaaac	364
Db	456	TCGTGAAGAACATAAAGTGGCAGACACTGCAGGCTGTAAATTTCTGCCATAAACATAACT	515
QY	365	gtattcacagagataataaacctgaaaaatattcttaataaagcaaggaataataacaaga	424
Db	516	GCATACACAGAGACGCGTGAAGCCCGGAAAAATATTCTCATCACCAAACAGTCAGCCATT	575
QY	425	tttgtgacttcgggtttgtgcacaaaattctgatt---ccaggagatgcctacacccgattatg	481
Db	576	TCTGTGACTTTGGGTCGACGCGCTCCTCACTGGACCANGTCACTACTACACAGACTACN	635
QY	482	tagctacgagatggtaccgagctcctgaaacttcttgtggagagatactcagtatggttctt	541
Db	636	GTGCCCCACCGGTGACTGCTCACCCGAGCTGCTAGTTGGAGACACCGCAGTATTGGTCCC	695
QY	542	cagtcgatatatgggctattggtgtgtgttt	572
Db	696	CTGTAGATGTCTGGGCAATTGGCTGTGTGT	726

RESULT 5

AA286088

LOCUS

DEFINITION

AA286088 718 bp mRNA EST 09-APR-1997
vc33b05.r1 Barstead MPLRB1 Mus musculus cDNA clone IMAGE:776337 5' similar to gb:X66358 SERINE/THREONINE-PROTEIN KINASE KKIALRE (HUMAN); mRNA sequence.

ACCESSION

AA286088

VERSION

AA286088.1 GI:1932198

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
The WashU-HHMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:469193

TITLE

JOURNAL

COMMENT

FEATURES


```
source
1. .718
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/clone_lib="Barstead MPLRB1"
/sex="mixed"
/tissue_type="kidney"
/dev_stage="6 weeks"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTACGAATCTGAAGTGGGAGCGGCCCTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[CATGGATTCGGTACC], digested with Not I and cloned into the
Not I and Eco RI sites of the modified pT7T3 vector.
Library constructed by Bob Barstead."
BASE COUNT 199 a 171 c 186 g 162 t
ORIGIN

Query Match 24.5%; Score 255; DB 10; Length 718;
Best Local Similarity 67.9%; Pred. No. 3.8e-47;
Matches 416; Conservative 0; Mismatches 190; Indels 7; Gaps 4;

QY 1 atggaagaagtataaaaaattagctaagactggagaagggtccttatggggttgattcaaa 60
Db 109 ATGGAATAATGAAAAAATGGAAGATTGGAGAAGGCTCCTATGGGTAGTGTTCAG 168

QY 61 tgcagaaacaaaacctctgtgacaaagtagtagctgttaaaaaatttgtggaatctgaagat 120
Db 169 TGCAGAAACAGGGACACCGGTCAGATCTGTGGCCATCAAGAGGTTTCTGGAAACCGAAGAT 228

QY 121 gatcctgtgttaagaaaaatagcactaagagaaaatacgtatgttgaagcaattaaacat 180
Db 229 GACCCCTGTCATAAAGAAATCGCCCTTCGAGAAATCGCATGCTCAAGCAACTCAAGCAC 288

QY 181 ccaaatcttgtgaacctcatcgaggtgttcaggagaaaaaggaaatgcattagtttt 240
Db 289 CCCAACCTGGTCAACCTCTCTGGAAGTCTTCCGGAG-AAGCGGAGCTTCACCTGGTGTTC 347

QY 241 gaatacttgtatcacatacacttttaaatgagctggagaagaaacccaaatggagtgtgat 300
Db 348 GAGTACTGCGACCACACACGGTGTCTCACGAGCTGGATCGGTA-TCAGAGGGGGGTACAGAG 406

QY 301 ggaagtataaaagcgtattatggcaaacacttcaagctcttaattctgtcatatacat 360
Db 407 CCTCTCGTGAAGAACATAACTTTGGCAGACACTGCAGGCTGTAAATTTCTGCCATAAACAT 466

QY 361 aactgtattcacagagatataaaacctgaaaaatattctaataactaagcaaggaataatc 420
Db 467 AACTGTCATACAGAGACGCTGAAGCCGGAAAAATATTCTCATTTACCAACAAGTCAGCCATT 526

QY 421 aagatttgtgacttcgggttttgc-----caaatcttgattccaggagatgcctacaccga 476
Db 527 AAGCTCTGTGACTTTGGGGTTCGAACGGCTTCTCACTGGACCCAGGTGACTACTACACGGA 586

QY 477 ttatgtagctacgagatggtaccgagctcctgaacttcttgtgggagatactcagtatgg 536
Db 587 CTACGTGGCCACCCGGTGGTACCCTCAACCGAGCTGCTAGTGGGAGACACCGCAGTATGG 646

QY 537 ttcttcagtcgatataatgggctattgggtgtgttttttgagagctcctgacagggccagcc 596
Db 647 TCCCCCTGTAGATGCTCTGGGC-AAATGGCTGTGTGTTTGTCTGAGTTACTGTCCGGAGTGCC 705

QY 597 actgtggcctgga 609
Db 706 TCTATGGCAAGGA 718

RESULT 6
AW233105
```

```
LOCUS AW233105 579 bp mRNA EST 30-JUL-2001
DEFINITION fj28b10.y1 zebrafish adult olfactory Danio rerio cDNA 5' similar to
gb:X66358 SERINE/THREONINE-PROTEIN KINASE KIAI2RE (HUMAN);, mRNA
sequence.
ACCESSION AW233105
VERSION AW233105.1 GI:6565408
KEYWORDS EST.
SOURCE zebrafish.
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Rasboridae; Danio.
REFERENCE 1 (bases 1 to 579)
AUTHORS Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
and Wilson,R.
TITLE WashU Zebrafish EST Project 1998
JOURNAL Unpublished (1998)
COMMENT Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrafish@watson.wustl.edu
cDNA Library Preparation: John Ngai cDNA Library Arrayed by:
Matthew Clark. DNA Sequencing by: Washington University Genome
Sequencing Center Clone Distribution: Genome Systems, St. Louis,
Missouri (web address: www.genomesystems.com) (email contact:
info@genomesystems.com) and Research Genetics, Huntsville, Alabama
(web address: www.resgen.com) (email contact: info@resgen.com) and
RessourcenZentrumPrimarDatenbank, Berlin, Germany (web address:
www.rzpd.de)
Seq primer: T3 ET from Amersham
High quality sequence stop: 404.
FEATURES
Location/Qualifiers
1. .579
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone_lib="zebrafish adult olfactory"
/sex="mixed"
/tissue_type="Olfactory rosettes"
/dev_stage="adult"
/lab_host="D10Hb (Gibco BRL)"
/note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI; This
is a directionally cloned cDNA library from adult
zebrafish olfactory epithelium."
BASE COUNT 174 a 110 c 137 g 158 t
ORIGIN

Query Match 20.2%; Score 209.8; DB 10; Length 579;
Best Local Similarity 73.4%; Pred. No. 5.1e-37;
Matches 268; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 1 atggaaaagtatgaaaaattagctaagactggagaagggtccttatggggttgattcaaa 60
Db 215 ATGGAGAAGTATGAGAAGATCAGTAAGATTGGCGAAGGCTCGTACGGCCTTGTGTTCAAG 274

QY 61 tgcagaaacaaaacctctgtgacaagtagtagctgttaaaaaatttgtggaatctgaagat 120
Db 275 TGCAGGAATAAAGACACCGGACAGATTGTCGCCATCAGAGAATTCGCCATCTGAAGCAACTAAACAT 334

QY 121 gatcctgtgttaagaaaaatagcactaagagaaaatacgtatgttgaagcaattaaacat 180
Db 335 GATCCCATCATTAAGAAAAATAGCGCTCAGAGAAATCCGCATCTCTGAAGCAACTAAACAT 394

QY 181 ccaaatcttgtgaacctcatcgaggtgttccaggagaaaaaggaaatgcatttagtttt 240
Db 395 CCAAACCTGGTGAATCTGTATGGAGGTGTTTCAGAGAAAGAGAAAAACTTCACCTGGTGT 454
```


QY	241	gaatactgtgatcatacaactttttaaatgagctgggaaagaaaccccaaatggagttgctgat	300
Db	455	GAGTGTGAGACCACTACTGTGCTGAATGAGCTGGACAGATACCCACGAGGTGTCCAGAG	514
QY	301	ggagtgatacaaaagcgtattatggcaaacacttcaagctctottaattctgtcatatacat	360
Db	515	CATATGGTTAAAGCATCATCTGGGCAACACTTTAGGTGTGAACTTCTGCCACAACAA	574
QY	361	aactg 365	
Db	575	AATTG 579	
RESULT	7		
AA626859/c			
LOCUS	AA626859	295 bp mRNA	EST 15-OCT-1997
DEFINITION	zu89f10.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:745195		
	3', similar to SW:KKIA_HUMAN Q00532 SERINE/THREONINE-PROTEIN KINASE		
	KKIALRE ;, mRNA sequence.		
ACCESSION	AA626859		
VERSION	AA626859.1	GI:2539246	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	1 (bases 1 to 295)		
	Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,		
	Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin		
	,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,		
	White,Y., Wylie,T., Waterston,R. and Wilson,R.		
TITLE	WashU-NCI human EST Project		
JOURNAL	Unpublished (1997)		
COMMENT	Contact: Wilson RK		
	Washington University School of Medicine		
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108		
	Tel: 314 286 1800		
	Fax: 314 286 1810		
	Email: est@watson.wustl.edu		
	This clone is available royalty-free through LNL ; contact the		
	IMAGE Consortium (info@image.llnl.gov) for further information.		
	Seq primer: -40ml3 fwd. ET from Amersham		
	High quality sequence stop: 292.		
FEATURES	Location/Qualifiers		
source	1. .295		
	/organism="Homo sapiens"		
	/db_xref="GDB:5933055"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:745195"		
	/clone_lib="Soares_testis_NHT"		
	/sex="male"		
	/lab_host="DH10B"		
	/note="Vector: pT7T3D-Pac (Pharmacia) with a modified		
	polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA		
	was prepared from mRNA obtained from Clontech Laboratories		
	, Inc., and primed with a Not I - oligo(dT) primer [5',		
	TGTTACCAATCTGAAGTGGAGCGCGCCCAATTGTTTTTT 3'].		
	Double-stranded cDNA was ligated to Eco RI adaptors		
	(Pharmacia), digested with Not I and cloned into the Not I		
	and Eco RI sites of the modified pT7T3 vector. Library		
	went through one round of normalization to Cot5, and was		
	constructed by Bento Soares and M. Fatima Bonaldo. "		
BASE COUNT	95 a 69 c 55 g 76 t		
ORIGIN			
Query Match	19.1%;	Score 199;	DB 10; Length 295;
Best Local Similarity	95.3%;	Pred. No. 1.4e-34;	
Matches	205; Conservative	0; Mismatches	10; Indels 0; Gaps 0;
QY	438	gtttgcacaaattctgattccaggagatgctacaccgattatgtagctacgaatggta	497
Db	280	GATTGATGCATTTCAGATTCCAGGAGATGCCTACACCATATGTAGCTACGAGATGGTA	221

QY	498	ccgagctcctgaactctttgtgggagatactcagtatggttcttcagtcgatatatgggc	557
Db	220	CCGAGCTCCTGAACCTCTTGTGGAGATACTCAGTATGGTTCTTCAGTCGATATATGGGC	161
QY	558	tattggttggttttgcagagctcctgcagagccagccactgtggcctggaataatcaga	617
Db	160	TATTGTTGTGTTTTTGCAGAGCTCCTGCAGGCCAGCCACTGTGGCCTGGAATAATCAGA	101
QY	618	tgtggaccaacttttatctgataaatcagaacactag	652
Db	100	TGTGGACCAACTTTATCTGTGATAATCAGAACACTAG	66
RESULT	8		
AI508835			
LOCUS	AI508835	727 bp mRNA	EST 12-MAR-1999
DEFINITION	vc33b05.y1 Barstead MPLRB1 Mus musculus cDNA clone IMAGE:776337 5'		
	similar to gb:X66358 SERINE/THREONINE-PROTEIN KINASE KKIALRE (HUMAN		
);, mRNA sequence.		
ACCESSION	AI508835		
VERSION	AI508835.1	GI:4407740	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	1 (bases 1 to 727)		
	Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,		
	Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person		
	,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter		
	,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,		
	Waterston,R. and Wilson,R.		
TITLE	The WashU-NCI Mouse EST Project 1999		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Marra M/WashU-NCI Mouse EST Project 1999		
	Washington University School of Medicine		
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA		
	Tel: 314 286 1800		
	Fax: 314 286 1810		
	Email: mouseest@watson.wustl.edu		
	This clone is available royalty-free through LNL ; contact the		
	IMAGE Consortium (info@image.llnl.gov) for further information.		
	MGI:469193		
	This read is a RESEQUENCE of a previously sequenced mouse clone		
	This read has been verified (found to hit its original self in the		
	correct orientation)		
	Seq primer: -40RP from Gibco		
	High quality sequence stop: 399.		
FEATURES	Location/Qualifiers		
source	1. .727		
	/organism="Mus musculus"		
	/strain="BALB/c"		
	/db_xref="taxon:10090"		
	/clone="IMAGE:776337"		
	/clone_lib="Barstead MPLRB1"		
	/sex="mixed"		
	/tissue_type="Kidney"		
	/dev_stage="6 weeks"		
	/lab_host="DH10B"		
	/note="Vector: pT7T3D-Pac (Pharmacia) with a modified		
	polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA		
	was primed with a Not I - oligo(dT) primer [5',		
	TGTTACGAATCTGAAGTGGAGCGCGCCCTTTT 3']		
	3']; double-stranded cDNA was ligated to Eco RI adaptors		
	[CATGGATTCGGTACC], digested with Not I and cloned into the		
	Not I and Eco RI sites of the modified pT7T3 vector.		
	Library constructed by Bob Barstead."		
BASE COUNT	200 a 169 c 191 g 158 t		
ORIGIN			
Query Match	17.9%;	Score 186.6;	DB 10; Length 727;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 14:28:48 ; Search time 116.84 Seconds
(without alignments)
2017.833 Million cell updates/sec

Title: US-09-671-050-9
Perfect score: 1041
Sequence: 1 atggaaaagtatgaaaaatt.....aggctacttcgcgtcaaaaagt 1041

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query %			DB ID	Description
	Score	Match	Length		
1	163.8	15.7	903	2	US-08-874-347-9
2	163.8	15.7	903	3	US-09-093-522-9
3	162	15.6	1476	2	US-08-969-106-1
4	133.2	12.8	1089	1	US-08-154-915-1
5	133.2	12.8	1089	2	US-08-464-517-37
6	133.2	12.8	1089	2	US-08-246-361A-37
7	133.2	12.8	1089	3	US-08-463-772-37
8	133.2	12.8	1089	5	PCT-US93-09945-1
9	128	12.3	1002	1	US-08-463-090B-3
10	116	11.1	1070	1	US-08-463-090B-5
11	103	9.9	2747	2	US-08-874-347-1
12	103	9.9	2747	3	US-09-093-522-1
13	96.2	9.2	1673	4	US-09-347-801-15
14	87.4	8.4	6340	1	US-08-187-785-3
15	87.4	8.4	87350	3	US-08-781-891-79
16	87	8.4	8779	2	US-08-750-703-4
17	86.6	8.3	2624	1	US-08-032-382B-1
18	85.4	8.2	7216	2	US-08-750-703-3
19	84.4	8.1	688	6	5498694-3
20	84.2	8.1	1100	2	US-08-950-449A-18
21	84.2	8.1	1429	4	US-09-347-801-13
22	84.2	8.1	3774	2	US-08-950-449A-11
23	84.2	8.1	3813	1	US-08-469-421-11
24	84.2	8.1	3813	1	US-08-250-975-11
25	84.2	8.1	3813	2	US-08-605-002A-11
26	84.2	8.1	3813	5	PCT-US94-10529-11
27	84.2	8.1	4453	1	US-08-770-761A-4

28	84.2	8.1	4540	1	US-08-770-761A-6	Sequence 6, Appli
29	84.2	8.1	4621	1	US-08-770-761A-1	Sequence 1, Appli
c 30	84.2	8.1	246240	2	US-08-724-394A-20	Sequence 20, Appl
c 31	84.2	8.1	246240	2	US-08-724-394A-21	Sequence 21, Appl
c 32	84.2	8.1	246240	2	US-08-724-394A-22	Sequence 22, Appl
33	83.8	8.0	1200	1	US-08-011-398B-3	Sequence 3, Appli
34	83.8	8.0	1200	1	US-08-464-051-3	Sequence 3, Appli
35	83.8	8.0	1200	2	US-08-462-498-3	Sequence 3, Appli
36	83.6	8.0	7720	4	US-09-318-448-5	Sequence 5, Appli
c 37	83.4	8.0	35060	3	US-08-814-095-7	Sequence 7, Appli
38	83.2	8.0	912	2	US-08-627-610-13	Sequence 13, Appl
39	83.2	8.0	1424	1	US-08-403-634-3	Sequence 3, Appli
40	83.2	8.0	1424	4	US-08-913-441B-3	Sequence 3, Appli
41	82.8	8.0	2562	2	US-08-436-771-8	Sequence 8, Appli
42	82.8	8.0	2562	2	US-08-434-998-8	Sequence 8, Appli
43	82.8	8.0	2562	2	US-08-487-797-8	Sequence 8, Appli
44	82.8	8.0	2562	5	PCT-US95-02058-8	Sequence 8, Appli
c 45	82.8	8.0	176373	3	US-09-128-155-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1
US-08-874-347-9
; Sequence 9, Application US/08874347
; Patent No. 5863741
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; TITLE OF INVENTION: CARINII
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/874,347
; FILING DATE: 13-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/055001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070
; TELEFAX: 612-288-9696
; TELEX:
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 903 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...900
; OTHER INFORMATION:
US-08-874-347-9

```
Query Match      15.7%; Score 163.8; DB 2; Length 903;
Best Local Similarity 55.2%; Pred. No. 1.8e-36;
Matches 365; Conservative 0; Mismatches 287; Indels 9; Gaps 2;

Qy 1 atggaaaagtatgaaaaattagctaagactgagaaaggggtcttattgagggttatttcaaa 60
    |||| | || | |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 1 ATGGAGCAATATCAGAGGTTAGAGAAGATTGGAGAGGAACCTTATGGAGTTGTTATAAA 60

Qy 61 tgcagaaacaaacacctctggaacaagtagctgtttaaaaaattgtggaatctgaagat 120
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 GCAAAGGATCTTGAAAGTGGTACAATTGTAGCTCTTAAGAAATCCGATTAGAAGCAGAA 120

Qy 121 gatcctgtgttaagaaaaatagcactaagagaaaaatcgtatgttgaagcaattaaaaacat 180
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | |||
Db 121 GATGAGGGAGTTCCTAGTACAGCAATTCTGTGAGATATCATTGTGAAAGAGATGCACAAT 180

Qy 181 ccaaatcttgtgaacctcatcgagggtgttcaggagaaaaaggaaaaatgcatttagtttt 240
    ||| |||| | | | | | | | | | | | | | | | | | | | | | |
Db 181 GATAATGTTGTAGACTTTTGAATATAATTTCATCAAGAGTCACGTTTATATCTTGTTTT 240

Qy 241 gaatactgtgatcacacacttttaaatgagctggaagaaacccaaatg-----gagtt 294
    |||| |||| | | | | | | | | | | | | | | | | | | | | | |
Db 241 GAATTTCTTGATCTTGATTTAAAAAAGTATATGAATAGTATTCCAAGGACATGATGCTT 300

Qy 295 gctgatggagtatcaaaagcgtattatggcaaacacitcaagctcttaatttctgtcat 354
    | | | | ||| |||| | | | | | | | | | | | | | | | | | | |
Db 301 GGTGCAGAAATGATTAAGTTTATGTCACAACTTGTATCAGGTGTAAATATTTGTCAT 360

Qy 355 atacataactgtattcacagagatatataaacctgaaaaatatttctaataactaagcaagga 414
    || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 361 TCTCATCGTATCTTCATCGTACTTGAACCCACAAAATCTCTTATAGATCGAGAGGA 420

Qy 415 ataataagatttgtgacttcgggtttgacaaa---ttctgattccaggagatgcctac 471
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 421 AATCTTAAATTAGCAGATTTTGGCTTGCAGGCGGTTTGGTGTTCATTCGTTGGTTAT 480

Qy 472 accgattatgtactacgagatggtagccagctctgaacttcttgtggagatactcag 531
    || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 481 ACTCATGAAGTTGTTACACTTTGGTATCGTCTCCAGAAAGTCTTTTAGTGGTCCGACAA 540

Qy 532 tatggttcttcagtcgatatatgggctattggtgtgttttcgagagctcctgcagggc 591
    |||| | | | | | | | | | | | | | | | | | | | | | | | |
Db 541 TATGCAACAGCGCTTGATATATGGAGCATTTGGATGTATTTTGCAGAAATGGCTACAAA 600

Qy 592 cagccactgtggcctggaanaatcagatgtggaccaactttatctgataaatcagaaacacta 651
    |||| | | | | | | | | | | | | | | | | | | | | | | | |
Db 601 AAGCCATTATTCCAGGTGATTCTGAAATTGTGAAATATTAGAAATATTAGAAATATTA 660

Qy 652 g 652
Db 661 G 661

RESULT 2
US-09-093-522-9
; Sequence 9, Application US/09093522
; Patent No. 6015700
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; TITLE OF INVENTION: CARINII
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
```

```
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/093,522
; FILING DATE: 08-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/874,347
; FILING DATE: 13-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/055002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070
; TELEFAX: 612-288-9696
; TELEX:
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 903 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...900
; OTHER INFORMATION:
; US-09-093-522-9
```

```
Query Match      15.7%; Score 163.8; DB 3; Length 903;
Best Local Similarity 55.2%; Pred. No. 1.8e-36;
Matches 365; Conservative 0; Mismatches 287; Indels 9; Gaps 2;

Qy 1 atggaaaagtatgaaaaattagctaagactggaagaagggtcttattgagggttatttcaaa 60
    |||| | || | |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 1 ATGGAGCAATATCAGAGGTTAGAGAAGATTGGAGAGGAACCTTATGGAGTTGTTATAAA 60

Qy 61 tgcagaaacaaacacctctggaacaagtagctgtttaaaaaattgtggaatctgaagat 120
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 GCAAAGGATCTTGAAAGTGGTACAATTGTAGCTCTTAAGAAATCCGATTAGAAGCAGAA 120

Qy 121 gatcctgtgttaagaaaaatagcactaagagaaaaatcgtatgttgaagcaattaaaaacat 180
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | |||
Db 121 GATGAGGGAGTTCCTAGTACAGCAATTCTGTGAGATATCACTTTTGAAGAGATGCACAAT 180

Qy 181 ccaaatcttgtgaacctcatcgagggtgttcaggagaaaaaggaaaaatgcatttagtttt 240
    ||| |||| | | | | | | | | | | | | | | | | | | | | | |
Db 181 GATAATGTTGTAGACTTTTGAATATAAATTCATCAAGAGTCACGTTTATATCTTGTTTT 240

Qy 241 gaatactgtgatcacacacttttaaatgagctggaagaaacccaaatg-----gagtt 294
    |||| |||| | | | | | | | | | | | | | | | | | | | | | |
Db 241 GAATTTCTTGATCTTGATTTAAAAAAGTATATGAATAGTATTCCAAGGACATGATGCTT 300

Qy 295 gctgatggagtatcaaaagcgtattatggcaaacacitcaagctcttaatttctgtcat 354
    | | | | ||| |||| | | | | | | | | | | | | | | | | | | |
Db 301 GGTGCAGAAATGATTAAGTTTATGTCACAACTTGTATCAGGTGTAAATATTTGTCAT 360

Qy 355 atacataactgtattcacagagatatataaacctgaaaaatatttctaataactaagcaagga 414
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | |||
Db 361 TCTCATCGTATCTTCATCGTACTTGAACCCACAAAATCTCTTATAGATCGAGAGGA 420

Qy 415 ataataagatttgtgacttcgggtttgacaaa---ttctgattccaggagatgcctac 471
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 421 AATCTTAAATTAGCAGATTTTGGCTTGCAGGCGGTTTGGTGTTCATTCGTTGGTTAT 480

Qy 472 accgattatgtactacgagatggtagccagctctgaacttcttgtggagatactcag 531
    || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 481 ACTCATGAAGTTGTTACACTTTGGTATCGTCTCCAGAAAGTCTTTTAGTGGTCCGACAA 540

Qy 532 tatggttcttcagtcgatatatgggctattggtgtgttttcgagagctcctgcagggc 591
    |||| | | | | | | | | | | | | | | | | | | | | | | | |
Db 541 TATGCAACAGCGCTTGATATATGGAGCATTTGGATGTATTTTGCAGAAATGGCTACAAA 600

Qy 592 cagccactgtggcctggaanaatcagatgtggaccaactttatctgataaatcagaaacacta 651
    |||| | | | | | | | | | | | | | | | | | | | | | | | |
Db 601 AAGCCATTATTCCAGGTGATTCTGAAATTGTGAAATATTAGAAATATTAGAAATATTA 660

Qy 652 g 652
Db 661 G 661
```


;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,772
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-004C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1089 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 13..888
; US-08-463-772-37

Query Match 12.8%; Score 133.2; DB 3; Length 1089;
Best Local Similarity 52.3%; Pred. No. 6.5e-28;
Matches 344; Conservative 0; Mismatches 308; Indels 6; Gaps 2;

QY 1 atggaaaagtatgaaaaattagctaagactggagaagggtcttatggggttgattcaaa 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 13 ATGCAGAAATACGAGAAACTGGAAAGATTGGGAAGGCACCTACGGAACCTGTGTTCAAG 72

QY 61 tgcagaaacaaaacctctggacaagtagtagctgtttaaaaaattgtgaaatctgaagat 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 73 GCCAAAAACCGGGAGACTCATGAGATCGTGGCTCTAAAAACGGGTGAGGCTGGATGACGAT 132

QY 121 gatacctgtgttaagaaaaatagcactaagagaaatacgtatgttgaagcaattaaacat 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 133 GATGAGGGTGTGCCGAGTTCGCCCTCCGGGAGATCTGCCTACTCAAGGAGCTGAAGCAC 192

QY 181 ccaaatcttgtgaacctcatcgaggtgttcagagaaaaaggaaaaatgcaatttagtttt 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 193 AAGAACATCGTCAGGCTTCATGACGTCCTGCACAGCGCACAGAAAGCTGACTTTGGTTTT 252

QY 241 gaatactgtgatcatcacacttttaaatgagctggaagaaaccccaatggagttgctgat 300
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 253 GAATTCTGTGACCGAGGACCTGAAGCCCCCAGAACCTGCTAATAACAGGAATGGGAGCTG 312

QY 301 ggaagtgatcaaaaagcgtattatggcaaacacttcaagctcttaatttctgtcatatacat 360
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 313 GAGATTGTAAAGTCATTCTCTTCAGCTACTAAAAGGGCTGGGATTCTGTCTATAGCCCG 372

QY 361 aactgtattcacagagatatataaacctgaaatatcttaataactaagcaaggaataatc 420
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 373 AATGTGCTACACAGGGACCTGAAGCCCCCAGAACCTGCTAATAACAGGAATGGGAGCTG 432

QY 421 aagatttgtgacttcgggttttgacaaaa---ttctgattocaggagatgacctacaccgat 477
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 433 AAATTGGCTGATTTTGGCCTGGCTCGAGCCCTTTGGGATTCCCGTTCGCTGTTACTCAGCT 492

QY 478 tatgtagctacgagatggttacccgagctccgaactctctgtgggagatactcagtatggt 537
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 493 GAGGTGGTCACACTGTGGTACCCGCCACCGGATGTCCTCTTTGGGGCCCAAGCTGTACTCC 552

QY 538 tcttcagtcgatatatgggctatgttggtgtgtttttgcagagct---cctgacaggccag 594
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 553 ACGTCCATCGACATGTGTCAGCGCGCTGCATCTTTGCAGAGCTGGCCAAATGCTGGCGG 612

QY 595 ccactgtggcctggaaaaatcacagatgtgagccaactttatctgataaatcagaacactag 652
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 613 CCTCTTTTCCCGCAATGATGTCGATGACCAAGTTGAAGAGGATCTTCCGACTGCTGG 670

RESULT 8
PCT-US93-09945-1
; Sequence 1, Application PC/TUS9309945
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Cyclin Complex Rearrangement and Uses Related
; TITLE OF INVENTION: Thereto
; NUMBER OF SEQUENCES: 4
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/09945
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/991,997
; FILING DATE: 17-DEC-1992
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1089 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 13..888
; PCT-US93-09945-1

Query Match 12.8%; Score 133.2; DB 5; Length 1089;
Best Local Similarity 52.3%; Pred. No. 6.5e-28;
Matches 344; Conservative 0; Mismatches 308; Indels 6; Gaps 2;

QY 1 atggaaaagtatgaaaaattagctaagactggagaagggtcttatggggttgattcaaa 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 13 ATGCAGAAATACGAGAAACTGGAAAGATTGGGAAGGCACCTACGGAACCTGTGTTCAAG 72

QY 61 tgcagaaacaaaacctctggacaagtagtagctgtttaaaaaatttgtgaaatctgaagat 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 73 GCCAAAAACCGGGAGACTCATGAGATCGTGGCTCTAAAAACGGGTGAGGCTGGATGACGAT 132

QY 121 gatacctgtgttaagaaaaatagcactaagagaaatacgtatgttgaagcaattaaacat 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 133 GATGAGGGTGTGCCGAGTTCGCCCTCCGGGAGATCTGCCTACTCAAGGAGCTGAAGCAC 192

QY 181 ccaaatcttgtgaacctcatcgaggtgttcagagaaaaaggaaaaatgcaatttagtttt 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 193 AAGAACATCGTCAGGCTTCATGACGTCCTGCACAGCGCACAGAAAGCTGACTTTGGTTTT 252

QY 241 gaatactgtgatcatcacacttttaaatgagctggaagaaaccccaatggagttgctgat 300
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 253 GAATTCTGTGACCGAGGACCTGAAGAAAGTATTTTGACAGTTGCAATGGTGACCTCGATCCT 312

QY 301 ggaagtgatcaaaaagcgtattatggcaaacacttcaagctcttaatttctgtcatatacat 360
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 313 GAGATTGTAAAGTCATTCTCTTCCAGCTACTAAAAGGGCTGGGATTCTGTCTATAGCCCG 372

QY 353 atatacataactgtattccacagagatataaaacctgaaaatatcttctaataactaagcaag 412
||| ||| | |||| | || | || | |||| |||| | | | |||
Db 1411 ATTCTCATCGTATTCTTCATCGTGACTTGAAGCCCCAAAATCTTCTTATTGATCGAGAAG 1470
QY 413 gaataatacaagatttgtgacttcgggtttgcac---aaattctgattccaggagatgcct 469
||| | ||| || |||| |||| |||| | || | |||| |||| | || |
Db 1471 GAAATCTTAAACTTGTCTGATTTCGGGCTTGCTCGGGCATTTGGTGTTCTTTACGTGGTT 1530
QY 470 acaccgattatgtagctacgagatggtaccgagctcctgaacttcctgtgtggagatactc 529
| || ||| || | || |||| || |||| || |||| || |||| | || |
Db 1531 ATACTCATGAGGTTGTACACTTTGGTATCGTGCTCCAGAAGTTCCTTTAGGTGGTCGAC 1590
QY 530 agtatggttcttcagtcogatatatgggctattggtgtgtttttgcagagctcctcgacag 589
| |||| | || | |||| || |||| || |||| || |||| || || |
Db 1591 AATATGCAACAGCACTTGATATTTGGAGCATTTGGATGTATTTTTCGCGAAATGGCTACAA 1650
QY 590 gccagccactgtggcctggaaaatcacagatgtggaccacactttatctgataataatcagaaca 648
| || | | || | || | || | || | || | || | || | || | || |
Db 1651 AAAAACCGTTATTTCCAGGCGATTCTGAAATTGATGAAATATTCAGAAATATTTTAGGTCA 1709

RESULT 12
US-09-093-522-1
; Sequence 1, Application US/09093522
; Patent No. 6015700
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; TITLE OF INVENTION: CARINII
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/093,522
; FILING DATE: 08-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/874,347
; FILING DATE: 13-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/055002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070
; TELEFAX: 612-288-9696
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2747 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 857...893
; OTHER INFORMATION: (A
; LOCATION: 942...1096

; OTHER INFORMATION:
; NAME/KEY: Coding Sequence
; LOCATION: 1147...1194
; OTHER INFORMATION:
; NAME/KEY: Coding Sequence
; LOCATION: 1242...1254
; OTHER INFORMATION:
; NAME/KEY: Coding Sequence
; LOCATION: 1311...1706
; OTHER INFORMATION:
; NAME/KEY: Coding Sequence
; LOCATION: 1751...1891
; OTHER INFORMATION:
; NAME/KEY: Coding Sequence
; LOCATION: 1944...2048
; OTHER INFORMATION:
US-09-093-522-1

Query Match 9.9%; Score 103; DB 3; Length 2747;
Best Local Similarity 57.4%; Pred. No. 2.6e-19;
Matches 206; Conservative 0; Mismatches 150; Indels 3; Gaps 1;

QY 293 ttgctgatggagtgatcaaaaagcgtattattatggcaaacactctcttaatttctgtc 352
||| | || | |||| |||| | || | || | |||| |||| | || |
Db 1351 TTGGAGCTGAAATGATCAAAAAATTTATGTCTCAACTTGTATCAGGTGTTAAATATTGTC 1410
QY 353 atatacataactgtattccacagagatataaaacctgaaaatatcttctaataactaagcaag 412
|| | || | |||| | || | || | |||| |||| | || |
Db 1411 ATTCTCAUCGTATTCTTCATCGTGACTTGAAGCCCCAAAATCTTCTTATTGATCGAGAAG 1470
QY 413 gaataatacaagatttgtgacttcgggttttgcaac---aaattctgattccaggagatgcct 469
||| | || | || |||| |||| |||| | || | |||| |||| | || |
Db 1471 GAAATCTTAAACTTGCCTGATTTCGGGCTTGCTCGGGCATTTGGTGTTCTTTACGTGGTT 1530
QY 470 acaccgattatgtagctacgagatggtaccgagctcctgaacttcctgtgtggagatactc 529
| || || | || | || |||| || |||| || |||| || |||| | || |
Db 1531 ATACTCATGAGGTTGTACACTTTGGTATCGTGCTCCAGAAGTTCCTTTAGGTGGTCGAC 1590
QY 530 agtatggttcttcagtcogatatatgggctattggtgtgtttttgcagagctcctcgacag 589
| |||| | || | |||| || |||| || |||| || |||| || || |
Db 1591 AATATGCAACAGCACTTGATATTTGGAGCATTTGGATGTATTTTTCGCTGAAATGGCTACAA 1650
QY 590 gccagccactgtggcctggaaaatcacagatgtggaccacactttatctgataataatcagaaca 648
| || | | || | || | || | || | || | || | || | || | || |
Db 1651 AAAAACCGTTATTTCCAGGCGATTCTGAAATTGATGAAATATTCAGAAATATTTTAGGTCA 1709

RESULT 13
US-09-347-801-15
; Sequence 15, Application US/09347801
; Patent No. 6262345
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Lee, Jian Ming
; TITLE OF INVENTION: Plant Protein Kinases
; FILE REFERENCE: BB-1171
; CURRENT APPLICATION NUMBER: US/09/347,801
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: 60/092,438
; EARLIER FILING DATE: July 10, 1998
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 1673
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1349)
US-09-347-801-15

```

Query Match          9.2%; Score 96.2; DB 4; Length 1673;
Best Local Similarity 61.7%; Pred. No. 1.6e-17;
Matches 171; Conservative 0; Mismatches 103; Indels 3; Gaps 1;

Qy 370 cacagagatataaaacctgaaaaatattcttaataactaaagcaa---ggaataataccaagatt 426
      ||||| ||| | ||||| ||||| ||| | | | | | | | | | | | | | | | | | | | |
Db 599 cacagggatgtgaaaccacacaaaatgtttggtgatctcttaaccocatcaagtcaagatc 658

Qy 427 tgtgacttcgggttgacaaaattctgattccagaggagatgcctacacccgattatgtagct 486
      ||||| ||| | ||||| ||||| ||| | | | | | | | | | | | | | | | | | | |
Db 659 tgtgactttggaagtgcaaaagtctctgttacctggtgaacccaacatagcatacatatgc 718

Qy 487 acgagatggtaccgagctcctgaaacttcttgtgggagatactacgtatggttcttcagtc 546
      | | | | | ||||| ||| | | | | | | | | | | | | | | | | | | | | |
Db 719 tctcgctactatcgtgctcctgagtcacatatttggtgcaactgaatatacaaaacttcaata 778

Qy 547 gatataatgggctattggtgtgtttttgcagagctcctgcagagccaccactgtggcct 606
      | | ||||| | | | ||||| ||||| ||||| ||| | | | | | | | | | | | | |
Db 779 gacataatggtcagctggatgtgttcttgcagagctacttcttggtcagcctctgtttcca 838

Qy 607 gaaaaatcagatgtggaccaaactttatctctgataatca 643
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 839 ggagagactgcggtgatcagctagtgagagattatca 875

RESULT 14
US-08-187-785-3/C
; Sequence 3, Application US/08187785
; Patent No. 5756476
; GENERAL INFORMATION:
; APPLICANT: Epstein, Stephen
; APPLICANT: Unger, Ellis
; APPLICANT: Speir, Edith
; TITLE OF INVENTION: Inhibition of No. 5756476-Transformed Cell
; TITLE OF INVENTION: Proliferation Using Anti-Sense Oligonucleotides
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson, and Bear
; STREET: 620 Newport Center Dr. Sixteenth Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/187,785
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/821,415
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E.
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: NIH001.001A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 714-760-0404
; TELEFAX: 714-760-9502
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6340 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Human
```

```

; IMMEDIATE SOURCE:
; CLONE: PCNA Genomic Clone
US-08-187-785-3

Query Match          8.4%; Score 87.4; DB 1; Length 6340;
Best Local Similarity 75.2%; Pred. No. 8.7e-15;
Matches 109; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Qy 650 tagtaagacggggtttcgccoatgttgaccaggctggtctcgaactcttgcagctcaagt 709
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 236 TAGTAGAGACGGGGTTTCACCATGTTGGCAGGCTGCTTTGAACACCTGACCTCAGGTG 167

Qy 710 atccacctgcccgtagcctctcaaaagtgcctggaattacaggaaaaattaatcccaagacatc 769
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 166 ATCCACCTGCCTCGCCTCTCAAAGTGCTGGGATTACAGGCTTGAGCCACCCAGCCCGCCAGC 107

Qy 770 aatcaatctttaaaagtaaacgggtt 794
      | ||| ||| || ||||| |
Db 106 CAGAAATTTTCAATTGTTACAGGAT 82

RESULT 15
US-08-781-891-79
; Sequence 79, Application US/08781891
; Patent No. 6090620
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; APPLICANT: Yu, Chang-En
; APPLICANT: Oshima, Junko
; APPLICANT: Mulligan, John T.
; APPLICANT: Schellenberg, Gerald D.
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
; TITLE OF INVENTION: WERNER'S SYNDROME
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,891
; FILING DATE: 27-DEC-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: NO. 6090620tenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 240052.419
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 79:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 87350 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-781-891-79

Query Match          8.4%; Score 87.4; DB 3; Length 87350;
Best Local Similarity 89.5%; Pred. No. 3e-14;
Matches 94; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 648 actagtagagacgggggtttcgccoatgttgaccaggctggtctcgaactcttgacgtcaag 707
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

us-09-671-050-9.rni

Search completed: May 2, 2002, 17:09:45
Job time: 9657 sec

